

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:42:05 ; Search time 20.45 Seconds  
(without alignments)  
1746.089 Million cell updates

Title: US-09-328-877A-8  
Perfect score: 589  
Sequence: 1 MSKPKPTLNHGLVPVDLKS.....SRNGTEETKLLVLDFFDVK 589

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

```

Word size : 0
Total number of hits satisfying chosen parameters: 412676

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```

Database :
A_Geneseq_0601.*
1: /SIDSI/cgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDSI/cgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDSI/cgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDSI/cgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDSI/cgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDSI/cgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDSI/cgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDSI/cgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDSI/cgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDSI/cgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDSI/cgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDSI/cgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDSI/cgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDSI/cgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDSI/cgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDSI/cgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDSI/cgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDSI/cgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDSI/cgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDSI/cgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDSI/cgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDSI/cgdata/geneseq/geneseqp/AA2001.DAT.*

```

## SUMMARIES

Result No.	Score	Query		Length	DB ID	Description
		Match	%			
1	589	100.0	589	20	AAV22466	Human mammary sel-
2	589	100.0	589	22	AAB59198	Human mammary sel-
3	559	94.9	559	20	AAV22467	Human mammary sel-
4	559	94.9	559	22	AAB59199	Human mammary sel-
5	557	94.6	589	21	AAB01204	Human GTPase assoc
6	540	91.7	540	20	AAV22465	Human hippocampal
7	540	91.7	540	20	AAV22468	Human mammary sel-
8	540	91.7	540	22	AAB59197	Human hippocampal
9	540	91.7	540	22	AAB59200	Human mammary sel-
10	540	91.7	540	20	AAV22464	Human hippocampal
11	540	91.7	545	22	AAB59196	Human hippocampal

12	540	91.7	553	20	AAV22463	Human hippocampal
13	540	91.7	553	22	AAB59195	Human hippocampal
14	540	91.7	592	20	AAV22462	Human hippocampal
15	540	91.7	592	22	AAB59194	Human hippocampal
16	540	91.7	626	20	AAV22469	Human 6myc-N-sel-1
17	540	91.7	626	22	AAB59201	Protein encoded by
18	540	91.7	627	20	AAV22461	Human hippocampal
19	540	91.7	627	22	AAB59193	Human hippocampal
20	540	91.7	666	20	AAV22471	Human C-term mychi
21	540	91.7	666	22	AAB59203	C-terminal mychi
22	540	91.7	669	20	AAV22470	Human Cterm V5 his
23	540	91.7	669	22	AAB59202	C-terminal V5 his
24	35	5.9	122	21	AAAG01075	Human secreted pro
25	12	2.0	587	20	AAV03204	Amino acid sequenc
26	9	1.5	257	21	AAE27862	Protein fragment e
27	9	1.5	472	18	AAW23598	Human LYST2 polype
28	9	1.5	472	20	AAV32120	Human LYST-2 prote
29	9	1.5	703	18	AAW23599	Mouse LYST2 polype
30	9	1.5	789	20	AAV32131	Human LYST-2 prote
31	8	1.4	13	17	AAR91634	Alpha chemokine GR
32	8	1.4	14	17	AAR91631	Alpha chemokine GR
33	8	1.4	14	17	AAR91632	Alpha chemokine GR
34	8	1.4	15	17	AAR91623	Alpha chemokine GR
35	8	1.4	16	17	AAR91620	Alpha chemokine GR
36	8	1.4	16	17	AAR91621	Alpha chemokine GR
37	8	1.4	73	14	AAR36771	MIP-2alpha. Homo
38	8	1.4	73	14	AAR36772	MIP-2beta. Homo s
39	8	1.4	73	16	AAR66699	Human gro-beta che
40	8	1.4	73	16	AAR66700	Human gro-gamma ch
41	8	1.4	73	16	AAR66698	Human gro-alpha ch
42	8	1.4	73	17	AAR93394	Protein used to ge
43	8	1.4	73	18	AAW18024	Human chemokine gr
44	8	1.4	73	18	AAW18025	Human chemokine gr
45	8	1.4	73	18	AAW18026	Human chemokine gr

## ALIGNMENTS

RESULT 1  
 AAY22466  
 ID AAY22466 standard; Protein; 589 AA.  
 XX  
 AC  
 AAY22466;  
 XX  
 29-SEP-1999 (first entry)  
 XX  
 Human mammary sel-10 protein sequence.  
 XX  
 Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;  
 KW mammary gland therapy.  
 KW  
 OS  
 XX  
 Homo sapiens.  
 XX  
 WO9932623-A1.  
 XX  
 01-JUL-1999.  
 XX  
 17-DEC-1998; 98WO-US26820.  
 XX  
 19-DEC-1997; 97US-0068243.  
 PR  
 (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 Gurney ME, Li J, Pauley JM;  
 WPI: 1999-458026/38.  
 N-PSDB; AAX99702.  
 XX  
 New isolated human sel-10 polypeptides  
 Claim 24; Page 60-63; 91pp; English.  
 PS

CC This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human mammary sel-10 protein. The polypeptides  
CC can be used to alter presenilin function. Compounds which inhibit either  
CC the expression or the activity of the human sel-10 polypeptides may  
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and  
CC therefore may be useful for the prevention or treatment of Alzheimer's  
CC disease.  
XX  
XX  
SQ Sequence 589 AA;

Query Match 100.0%; Score 589; DB 20; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKPKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPFCRRMRKRLDHGSEV 60  
DB 1 mskpgkptlnhglvpdlksakeplphqtvnmkifsisiaaglpfcrmrkrkldhgs 60  
QY 61 RFSLSGKPKCVSEYTTGLVPCSAATPTFGDLRAANGQQRRTTSVOPPTGLQEWL 120  
DB 61 rfsalglkpkcvseytattglvpcsaatptfgdlraangqqrtritsvopptglqewl 120  
QY 121 KMFSWSGPEKLLALDELIDSCPTQVKHMMQVTEPOFQDFISLLPKELALYVLSLEP 180  
DB 121 kmfswsgpekllaldelidsceptqvkhumqvlepfqdfisllpkelalylvlslep 180  
QY 181 KDLQAATCRYWIRLAEDNLLREKKEGIDEPHLIKRRKVIKPGFIHSPKSAVIRQ 240  
DB 181 kdllqaatcrywirlaednllwrekkeegideplhikrrkvikpgfihspkwsayir 240  
QY 241 HRIDTNRRRGELKSPKVLKGHDHVTICLQFCGNRIYSGDDNTLKVMSAVTGKCLRTL 300  
DB 241 hridtnrrrgelkspkvlkgdhdhviticlqfcgnriygsddntlkvmsavtgkclrt 300  
QY 301 GHTCGVSSQMRDNIISGSTDRTLKVMNAETGECIHTLGHQSLTSGMELKDNILVSGN 360  
DB 301 ghtcgvvssqmrndniisgstdrtlkvmaetgecihtlghqsltsgmelkdnilvsg 360  
QY 361 RDTLRVWDIETGQCLHVLGMHVAARVCQYDGRVRVSGAYDFMVKVWDPEPETCLHTLQ 420  
DB 361 rdtlrvwdi etgqclhvlgmhvaarvcydgrrvrvsgaydfmrvkvw dpetetclhtl 420  
QY 421 GHTNRVYSLQFDGIHVYSGSLDTSIRVMDVETGNCIHTLGHQSLTSGMELKDNILVSGN 480  
DB 421 ghtnrvy slqfdgi hvysgsl dtsirvmdvetgncihtlghqsltsgmelkdnilvsg 480  
QY 481 ADSTVKIWDIKTGQCLQTLQGPKNHQSAVTCLOPNKFNVTSSDDGVKLDLKTGEFIR 540  
DB 481 adstvk iwdikt gqclt lqgpnkhgsavtclqfnknfvitssddgtvkl dktgefir 540  
QY 541 NLVTLESGGGVGVWRIRASNTKLVCAGSRNGTEETKLLVLDLFDVDMK 589  
DB 541 nlvtlesggggv vwrirasntklvcagsrngteetkllvldlfdvdmk 589

RESULT 2  
AAB59198  
ID AAB59198 standard; protein; 589 AA.

AC AAB59198  
XX  
XX  
DT 23-MAR-2001 (first entry)  
XX Human mammary sel-10-1 protein.  
XX Sel-10; human; Alzheimer's disease; Abeta.  
XX Homo sapiens.  
XX W0200075328-A1.  
XX  
XX 14-DEC-2000.

XX 23-MAY-2000; 2000WO-US09814.  
XX 09-JUN-1999; 99US-0328877.  
XX (PHAA) PHARMACIA & UPJOHN CO.  
XX Pauley AM, Gurney ME, Li J;  
XX WPI; 2001-102404/A1.  
XX  
XX New human sel-10 polypeptides and their encoding polynucleotides,  
XX useful for raising antibodies for detecting sel-10 polypeptide  
XX expression and as drug targets in the treatment of Alzheimer's disease  
XX  
XX Claim 1; Page 79-82; 116pp; English.  
XX  
XX The present invention relates to human sel-10. The sel-10 proteins of  
XX the invention are useful for raising monoclonal or polyclonal  
XX antibodies useful in diagnostic assays for detecting sel-10  
XX polypeptide expression. The sel-10 polypeptides are also useful as drug  
XX targets for decreasing antibody levels in the treatment of Alzheimer's  
XX disease. It is also useful for identifying agents capable of  
XX altering the production level of Abeta. The polynucleotides are useful  
XX for developing assays for identifying agents capable of interfering  
XX with the biological pathways that lead to Alzheimer's disease.  
XX  
XX Sequence 589 AA;

Query Match 100.0%; Score 589; DB 22; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKPKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPFCRRMRKRLDHGSEV 60  
DB 1 mskpgkptlnhglvpdlksakeplphqtvnmkifsisiaaglpfcrmrkrkldhgs 60  
QY 61 RFSLSGKPKCVSEYTTGLVPCSAATPTFGDLRAANGQQRRTTSVOPPTGLQEWL 120  
DB 61 rfsalglkpkcvseytattglvpcsaatptfgdlraangqqrtritsvopptglqewl 120  
QY 121 KMFSWSGPEKLLALDELIDSCPTQVKHMMQVTEPOFQDFISLLPKELALYVLSLEP 180  
DB 121 kmfswsgpekllaldelidsceptqvkhumqvlepfqdfisllpkelalylvlslep 180  
QY 181 KDLQAATCRYWIRLAEDNLLREKKEGIDEPHLIKRRKVIKPGFIHSPKSAVIRQ 240  
DB 181 kdllqaatcrywirlaednllwrekkeegideplhikrrkvikpgfihspkwsayir 240  
QY 241 HRIDTNRRRGELKSPKVLKGHDHVTICLQFCGNRIYSGDDNTLKVMSAVTGKCLRTL 300  
DB 241 hridtnrrrgelkspkvlkgdhdhviticlqfcgnriygsddntlkvmsavtgkclrt 300  
QY 301 GHTCGVSSQMRDNIISGSTDRTLKVMNAETGECIHTLGHQSLTSGMELKDNILVSGN 360  
DB 301 ghtcgvvssqmrndniisgstdrtlkvmaetgecihtlghqsltsgmelkdnilvsg 360  
QY 361 RDTLRVWDIETGQCLHVLGMHVAARVCQYDGRVRVSGAYDFMVKVWDPEPETCLHTLQ 420  
DB 361 rdtlrvwdi etgqclhvlgmhvaarvcydgrrvrvsgaydfmrvkvw dpetetclhtl 420  
QY 421 GHTNRVYSLQFDGIHVYSGSLDTSIRVMDVETGNCIHTLGHQSLTSGMELKDNILVSGN 480  
DB 421 ghtnrvy slqfdgi hvysgsl dtsirvmdvetgncihtlghqsltsgmelkdnilvsg 480  
QY 481 ADSTVKIWDIKTGQCLQTLQGPKNHQSAVTCLOPNKFNVTSSDDGVKLDLKTGEFIR 540  
DB 481 adstvk iwdikt gqclt lqgpnkhgsavtclqfnknfvitssddgtvkl dktgefir 540  
QY 541 NLVTLESGGGVGVWRIRASNTKLVCAGSRNGTEETKLLVLDLFDVDMK 589  
DB 541 nlvtlesggggv vwrirasntklvcagsrngteetkllvldlfdvdmk 589

Db 541 nlvtlesggsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT 3

AAY22467  
ID AAY22467 standard; Protein; 559 AA.

XX

AC AAY22467;

XX 29-SEP-1999 (first entry)

XX

XX Human mammary sel-10 protein sequence.

XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;

KW mammary gland; therapy.

XX

XX Homo sapiens.

XX

PN W09932623-A1.

XX

PD 01-JUL-1999.

XX

PF 17-DEC-1998; 98WO-US26820.

XX

PR 19-DEC-1997; 97US-0068243.

XX

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX

PI Gurney ME, Li J, Pauley AM;

XX

DR WPI; 1999-458026/38.

XX

DR N-PSDB; AAX99702.

XX

XX New isolated human sel-10 polypeptides

PT

XX

PS Claim 24; Page 63-66; 91pp; English.

XX

CC This sequence represents a human sel-10 protein of the invention. This

CC sequence is specifically a human mammary sel-10 protein. The polypeptides

CC can be used to alter presenilin function. Compounds which inhibit either

CC the expression or the activity of the human sel-10 polypeptides may

CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and

CC therefore may be useful for the prevention or treatment of Alzheimer's

CC disease.

XX

XX

SQ Sequence 559 AA;

XX

Query Match 94.9%; Score 559; DB 20; Length 559;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 MKIFSISIIAQLPFCRRMRKLDHGSEVRSFSLGKPKCKVSEYTSWTGLVPCSATPTT 90

Db 1 mkifsisiiaqlpfcrrmrkldhgsevrfsfgkpkckvseysttstgylpcsatptt 60

QY 91 FGDRAANGQGOQRRRITSVQPTTGLQELKMFQSWSGPEKLLALDELIDSCPTQVKHM 150

Db 61 fgdlaangggqrrritsvqpttqlqewlkmfswsgpekllalidelidsceptqykhm 120

QY 151 MQVIEPQFORDISLLPKELALYVLSFLEPKDLQAQTCRYWRILAEONLLWREKCKEE 210

Db 121 mqviepqfdrdisllpkelalylvisflepkdlqlqaactcrywriilaednllwrekcke 180

QY 211 GIDPELHKRRKRVKPGFFTHSPKSNAYIQHRIDINWRREGELKSPKVLKGDHDDHVTCLQ 270

Db 181 gideplhkrkrvkvpgffthspksayirhridtnwrrgelkspkvlkgdhddhvtclq 240

QY 271 FCGNRIVSGSDNTLKVWSAVTGKCLRTLVGHTGGVSSQMRDNIITSGSTDRTLKVWNA 330

Db 241 fcgnrivsgsdntlkvwsvavgkclrtlvtgthggvssqmrndiitsgstdrtlkvwna 300

QY 331 ETGECIHTLYGHTSTVRCMHLHKEKRVVSGSRDATLRVNDIETGQCILHVLGMHVAARVCVQ 390

Db 301 etgecihtlyghtstvrcmhlhkekrvvsgrdatlrwvdiatgqclhvlmgvhaavrcvq 360

QY 391 YDGRRVVSGAYDFPMVKVNDPETETCLHTLQGTNRVYSLOFDGIHVVSGLDTSIRVWDV 450

Db 361 ydgrrvvsgaydfmrvkvwndpetetclhtlqgthnrvyqlqfdgihvvsghldtsirvwdv 420

QY 451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGOCLOTLOGPNKHOSAVT 510

Db 421 etgncihtltghsltsmelmknllvsgnadstvkiwdiktgcqcltldggnkhgsavt 480

QY 511 CLQFNKNEVITSSDDGTVKIWLDTGTEFIRNLVTLSESGSGVVRIRASNTKLVCAGVS 570

Db 481 clqfnknvitesddgtvkiwdiktgefirnlvtlesggsggvvwrirasntklvcavgs 540

QY 571 RGTETETKLLVLDVDMK 589

Db 541 rgtetetkllvldfdvdmk 559

XX

XX AAB59199

XX ID AAB59199 standard; protein; 559 AA.

XX AC AAB59199;

XX XX 23-MAR-2001 (first entry)

XX Human mammary sel-10-2 protein.

XX Sel-10; human; Alzheimer's disease; Abeta.

XX

XX Homo sapiens.

XX W0200075328-A1.

XX 14-DEC-2000.

XX 23-MAY-2000; 2000WO-US09814.

XX 09-JUN-1999; 99US-0328877.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX Pauley AM, Gurney ME, Li J;

XX WPI; 2001-102404/11.

XX New human sel-10 polypeptides and their encoding polynucleotides,

XX useful for raising antibodies for detecting sel-10 polypeptide

XX expression and as drug targets in the treatment of Alzheimer's disease

XX

XX Claim 1; Page 83-86; 116pp; English.

XX The present invention relates to human sel-10. The sel-10 proteins of

XX the invention are useful for raising monoclonal or polyclonal

XX antibodies useful in diagnostic assays for detecting sel-10

XX polypeptide expression. The sel-10 polypeptides are also useful as drug

XX targets for decreasing antibody levels in the treatment of Alzheimer's

XX disease. It is also useful for identifying agents capable of

XX altering the production level of Abeta. The polynucleotides are useful

XX for developing assays for identifying agents capable of interfering

XX with the biological pathways that lead to Alzheimer's disease.

XX

XX Sequence 559 AA;

SQ

Query Match 94.9%; Score 559; DB 22; Length 559;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 MKIFSISIIAQLPFCRRMRKLDHGSEVRSFSLGKPKCKVSEYTSWTGLVPCSATPTT 90





```
XX AAY22465;
AC 29-SEP-1999 (first entry)
DT Human hippocampal sel-10 protein sequence.
XX
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
KW therapy.
KW
XX Homo sapiens.
OS
XX W09932623-Al.
PN
XX 01-JUL-1999.
PD
XX 17-DEC-1998; 98WO-US26820.
XX
XX 19-DEC-1997; 97US-0068243.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Gurney ME, Li J, Pauley AM;
PI
XX WPI; 1999-458026/38.
DR
XX N-PSDB; AAX99701.
XX
XX New isolated human sel-10 polypeptides
PT
XX Claim 24; Page 57-59; 91pp; English.
PS
XX
XX This sequence represents a human sel-10 protein of the invention. This
CC sequence is specifically a human hippocampal sel-10 protein. The
CC polypeptides can be used to alter presenilin function. Compounds which
CC inhibit either the expression or the activity of the human sel-10
CC polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)
CC or PS-2, and therefore may be useful for the prevention or treatment of
CC Alzheimer's disease.
XX
XX Sequence 540 AA;

Query Match 91.7%; Score 540; DB 20; Length 540;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 MKRKLDHGEVRSFSLGKKPKCVSEYTSITGLVPCSATPTTFGLRAANGQQRRIITS 109
DB 1 mkrkidhgevrsfslgkpkcvseyststglvpcsatpttfgdraangqgrrrits 60

QY 110 VQPPGTGLQEWLKMFSWGSPEKLLALDELIDSCPTQVKHMQVIEPQFORDFISLLPKE 169
DB 61 vqpptglqewlkmfswgspekllalidelidsceptqvkhhmqviepqfgrdfisllpke 120

QY 170 LALYLSFLEPKDLQAACOTCRYWRILAEDNLLWREKCKEEDPELHLKRRKVIKPGFI 229
DB 121 lalylsflepkdlqaactcrywrlaednllwrekkeegideplhlkrrkvikpgfi 180

QY 230 HSPWKSAYIRQHRIDTNWRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVWS 289
DB 181 hspwksayirqhridtnwrgelkspkvlkgdhdhvttclqfcgnrivsgsdntlkvws 240

QY 290 AVTGKCLRTLVGHTGGVSSQMRDNIISGSTDRILKVMNAETGECIHTLYGHTSTVRGM 349
DB 241 avtgkclrtlvghtggvssqmrdniisgstdrilkvwnaetgecihtlyghtstvrvm 300

QY 350 HLHKEKRVSGSRDALTIRVWDIETGQCLHVLGMHVAARVCQYDGRVRVSGAYDFWVKVWD 409
DB 301 hlhkekrvsgsrdatlrvwdietgqclhvlgmhvaarvcydydgrvrsvgaydfmwkvwd 360

QY 410 PETETCLHTLOQHTNRVYSLOFGDTHVYSGSLDTSIRVWVDVTGNCIHTLTGHQSITSGM 469
DB 361 petetclhtloqhtnrvyslqfdghvsvsgsltdtsirvwvdvtgncihtltghqsltsgm 420
```

```
QY 470 ELKDNILVSGNADSTVKIWDIKTGQCLOTLQGNKKHQSAYTCLQFNKFVITSSDDGTVK 529
DB 421 elkdnilvsgnadstvkIwdiktgqcloTLqgnkqhqsavtcIqfnkfnvitsddgtvk 480

QY 530 LMDLKTGEFIRNLVLTLESGGGVWVRIRASNTKLCVACVSGSRGTETKLLVLDFFDVKM 589
DB 481 lwdlktgefirnlvltlesgggvwvrrasntklcvacvsgsrngteetklvlvldfdvdkm 540

RESULT 7
AAY22468
ID AAY22468 standard; Protein; 540 AA.
XX
XX AAY22468;
AC
XX 29-SEP-1999 (first entry)
DT Human mammary sel-10 protein sequence.
XX
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;
KW mammary gland; therapy.
KW
XX Homo sapiens.
OS
XX W09932623-Al.
PN
XX 01-JUL-1999.
PD
XX 17-DEC-1998; 98WO-US26820.
XX
XX 19-DEC-1997; 97US-0068243.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Gurney ME, Li J, Pauley AM;
PI
XX WPI; 1999-458026/38.
DR
XX N-PSDB; AAX99702.
XX
XX New isolated human sel-10 polypeptides
PT
XX Claim 24; Page 66-69; 91pp; English.
PS
XX
XX This sequence represents a human sel-10 protein of the invention. This
CC sequence is specifically a human mammary sel-10 protein. The polypeptides
CC can be used to alter presenilin function. Compounds which inhibit either
CC the expression or the activity of the human sel-10 polypeptides may
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and
CC therefore may be useful for the prevention or treatment of Alzheimer's
CC disease.
XX
XX Sequence 540 AA;
```

Query Match 91.7%; Score 540; DB 20; Length 540;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 50 MKRKLDHGEVRSFSLGKKPKCVSEYTSITGLVPCSATPTTFGLRAANGQQRRIITS 109
DB 1 mkrkidhgevrsfslgkpkcvseyststglvpcsatpttfgdraangqgrrrits 60

QY 110 VQPPGTGLQEWLKMFSWGSPEKLLALDELIDSCPTQVKHMQVIEPQFORDFISLLPKE 169
DB 61 vqpptglqewlkmfswgspekllalidelidsceptqvkhhmqviepqfgrdfisllpke 120

QY 170 LALYLSFLEPKDLQAACOTCRYWRILAEDNLLWREKCKEEDPELHLKRRKVIKPGFI 229
DB 121 lalylsflepkdlqaactcrywrlaednllwrekkeegideplhlkrrkvikpgfi 180

QY 230 HSPWKSAYIRQHRIDTNWRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVWS 289
DB 181 hspwksayirqhridtnwrgelkspkvlkgdhdhvttclqfcgnrivsgsdntlkvws 240
```

```
Db 181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrlysgsddntlkws 240
Qy 290 AVTGKCLRTFLVGHGTGGVWSSQMRDNIISGSTDRFLKVNNAETGBCIHTLYGHTSTVRGM 349
Db 241 avtgkclrtflvghgtggvssgmrdnliisgstdrflkvnnaetgbcihlyghtstvr 300
Qy 350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVQYDGRVVSAYDFWVKVWD 409
Db 301 hlhekrvvsgsrdatlrvwdietgqclhvlmgghvaavrcvqydgrrvvsaydfmfvkwd 360
Qy 410 PETETCLHTLQGHGHTNRVYSLOFDGTHVVSGLDTSIRVWDVETGNCIHTLGHGSLTSGM 469
Db 361 petetclhtlgghtnrvyslqfdghvvsghsldtsirvwdvetgncihltlghgsitsgm 420
Qy 470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTGK 529
Db 421 elkdnilysgnadstvkwi dktgqcltqgpnkhqsavtclqfnkfnvitsddgtvk 480
Qy 530 LWDLTGTFEIRNLVTLES GSGGVVWRIRASNTKLVCVAVSRNGTEETKLLVLDVFDVDMK 589
Db 481 lwdlktgfeirnlvtlesgsggvvwrirasntklvcavsrngteetkllvldfvdvdmk 540

RESULT 8
AAB59197
ID AAB59197 standard; protein; 540 AA.
AC AAB59197;
XX
XX
DT 23-MAR-2001 (first entry)
XX
XX Human hippocampal sel-10-5 protein.
DE
DE Sel-10; human; Alzheimer's disease; Abeta.
KW
OS Homo sapiens.
XX
XX WO2000075328-A1.
XX
XX 14-DEC-2000.
XX
XX 23-MAY-2000; 2000WO-US09814.
XX
XX 09-JUN-1999; 99US-0328877.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
PA
PI Pauley AM, Gurney ME, Li J;
XX
XX WPI; 2001-102404/11.
XX
XX New human sel-10 polypeptides and their encoding polynucleotides,
PT useful for raising antibodies for detecting sel-10 polypeptide
PT expression and as drug targets in the treatment of Alzheimer's disease
PT
XX
XX Claim 1; Page 75-78; 116pp; English.
XX
XX The present invention relates to human sel-10. The sel-10 proteins of
CC the invention are useful for raising monoclonal or polyclonal
CC antibodies useful in diagnostic assays for detecting sel-10
CC polypeptide expression. The sel-10 polypeptides are also useful as drug
CC targets for decreasing antibody levels in the treatment of Alzheimer's
CC disease. It is also useful for identifying agents capable of
CC altering the production level of Abeta. The polynucleotides are useful
CC for developing assays for identifying agents capable of interfering
CC with the biological pathways that lead to Alzheimer's disease.
XX
SQ Sequence 540 AA;
```

Query Match 91.7%; Score 540; DB 22; Length 540;  
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 50 MKRKLDHGEVSRSLSGKKPKCKVSEYTTTGLVPCSATPTTTFGDLRAANGOGQRRRITS 109
Db 1 mkrklhdgsevsrsfslgkpkckvseysttglvpcsatpttfgdlraangogqrrits 60
Qy 110 VQPTGTGLEWLKMQSGWSGPEKLLALDELIDSCETQVKHMQVIEPQFQRFISLLPKE 169
Db 61 vqptgtglewlkmgfswsgpekllalidelidsceptqvkhhmqviefpqfqrdfisllpke 120
Qy 170 LALVYLSLEPKDLLOAAQTCTRYNRILAEDNLLWREKCKEEDIDPLHKKRKVKIKPGFI 229
Db 121 lalyvlslepkdlloaaqtctrynrilaednllwrekckeegideplhkkrrkvikpgfi 180
Qy 230 HSPWKSAYIRQHRIIDTNWRRGELASPKVLKGHDDHVIITCLQFCGNRIYSGSDDNTLKWS 289
Db 181 hspwksayirqhridtnwrrgelkspkvlkghddhvitclqfcgnrlysgsddntlkws 240
Qy 290 AVTGKCLRTFLVGHGTGGVWSSQMRDNIISGSTDRFLKVNNAETGBCIHTLYGHTSTVRGM 349
Db 241 avtgkclrtflvghgtggvssgmrdnliisgstdrflkvnnaetgbcihlyghtstvr 300
Qy 350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVQYDGRVVSAYDFWVKVWD 409
Db 301 hlhekrvvsgsrdatlrvwdietgqclhvlmgghvaavrcvqydgrrvvsaydfmfvkwd 360
Qy 410 PETETCLHTLQGHGHTNRVYSLOFDGTHVVSGLDTSIRVWDVETGNCIHTLGHGSLTSGM 469
Db 361 petetclhtlgghtnrvyslqfdghvvsghsldtsirvwdvetgncihltlghgsitsgm 420
Qy 470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTGK 529
Db 421 elkdnilysgnadstvkwi dktgqcltqgpnkhqsavtclqfnkfnvitsddgtvk 480
Qy 530 LWDLTGTFEIRNLVTLES GSGGVVWRIRASNTKLVCVAVSRNGTEETKLLVLDVFDVDMK 589
Db 481 lwdlktgfeirnlvtlesgsggvvwrirasntklvcavsrngteetkllvldfvdvdmk 540

RESULT 9
AAB59200
ID AAB59200 standard; protein; 540 AA.
XX
XX AAB59200;
AC
XX
XX 23-MAR-2001 (first entry)
DT
XX
XX Human mammary sel-10-3 protein.
DE
XX
XX Sel-10; human; Alzheimer's disease; Abeta.
XX
XX Homo sapiens.
OS
XX
XX WO2000075328-A1.
PN
XX
XX 14-DEC-2000.
ED
XX
XX 23-MAY-2000; 2000WO-US09814.
PF
XX
XX 09-JUN-1999; 99US-0328877.
PR
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
PA
XX
XX Pauley AM, Gurney ME, Li J;
PI
XX
XX WPI; 2001-102404/11.
DR
XX
XX New human sel-10 polypeptides and their encoding polynucleotides,
PT useful for raising antibodies for detecting sel-10 polypeptide
PT expression and as drug targets in the treatment of Alzheimer's disease
PT
XX
XX Claim 1; Page 86-89; 116pp; English.
XX
XX
```

XX The present invention relates to human sel-10. The sel-10 proteins of  
 CC the invention are useful for raising monoclonal or polyclonal  
 CC antibodies useful in diagnostic assays for detecting sel-10  
 CC polypeptide expression. The sel-10 polypeptides are also useful as drug  
 CC targets for decreasing antibody levels in the treatment of Alzheimer's  
 CC disease. It is also useful for identifying agents capable of  
 CC altering the production level of Abeta. The polynucleotides are useful  
 CC for developing assays for identifying agents capable of interfering  
 CC with the biological pathways that lead to Alzheimer's disease.  
 XX  
 XX Sequence 540 AA;

Query Match 91.7%; Score 540; DB 22; Length 540;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 MRRKLDHGSEVRSFSLGKKPKCKVSEYTSITGLVPCSATPTTFGDLRAANGQGORRRITS 109  
 DB 1 mrrkldhgsevrslgkpkckvseysttglvpcsatpttfgdlraangqgqrrrits 60  
 QY 110 VQPPTGLQWLKMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQFQDFISLLPKE 169  
 DB 61 vqpptglqewlkmfqsqgpekllalidelidsceptqvkhhmqvleqpqfdrfislilpke 120  
 QY 170 LALYVLSFLEPKDLLQAATCRYWRILAEEDNLLWRECKEEDIDELPHIKRRKVIKPGFI 229  
 DB 121 lalyvlsflepkdllqaactcrywriilaednllwrekekeegideplhikrrkvikpgfi 180  
 QY 230 HSPWKSAIYRQHRIDNWRGELKSPKVLKGDHDDHVTITCLQFCGNRIVSGSDNTLKVWS 289  
 DB 181 hspwksaiyrqhridnwrngelkspkvlkgdhddhvtitclqfcgnrivsgsdntlkvws 240  
 QY 290 AVTGKCLRTLVTGHTGCVWSSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRGM 349  
 DB 241 avtgkclrtlvtghtgcvwssqmrndniisgstdrtlkvwnaetgecihtlyghtstvrGM 300  
 QY 350 HLHEKRVVSGSRDATLRVWMDIETGCQLHVLGMHVAARVCQYDGRVRSVSGAYDFMVKVWD 409  
 DB 301 hlhekrvvsgsrdatlrvmwdietgqclhvlgmhvaarvcydgrrrvsvsgaydfmfvkvwD 360  
 QY 410 PETETCLHTLQHTNRVYSLQFDGIHVVSGLDTSIRVWMDVETGNCIHTLTGHQSLTSGM 469  
 DB 361 petetclhtlqghtnrvyslqfdgihvvsldtsirvwdvetgncihtltghqsltsGM 420  
 QY 470 ELKDNILVSGNADSTVKIWDIKTGQCQLQTLQGNPKHQSATVCLQFNKNFVITSSDDGTVK 529  
 DB 421 elkdnilvsgnadstvkiwdiktgqclqltqgnpkhqsavtclqfnknfvitssddgtvk 480  
 QY 530 LWDLKTGEFIRNLVTLES GSGGVVWRIRASNTKLVCVSGSRNGTEETKLLVLDLDFVDMK 589  
 DB 481 lwdlktgefirnlvtlesgsggvvwrirasnktlvcavsgsrngteetkllvldfdvdmk 540

RESULT 10  
 AAY22464  
 ID AAY22464 standard; Protein; 545 AA.  
 XX  
 AC AAY22464;  
 XX  
 DT 29-SEP-1999 (first entry)  
 XX  
 DE Human hippocampal sel-10 protein sequence.  
 XX  
 KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO932623-A1.  
 XX  
 PD 01-JUL-1999.

XX 17-DEC-1998; 98WO-US26820.  
 XX 19-DEC-1997; 97US-0068243.  
 PR (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PA Gurney ME, Li J, Pauley AM;  
 PI WPI; 1999-458026/38.  
 XX DR N-PSDB; AAX99701.  
 XX  
 PT New isolated human sel-10 polypeptides  
 XX  
 PS Claim 24; Page 53-56; 91pp; English.  
 XX  
 CC This sequence represents a human sel-10 protein of the invention. This  
 CC sequence is specifically a human hippocampal sel-10 protein. The  
 CC polypeptides can be used to alter presenilin function. Compounds which  
 CC inhibit either the expression or the activity of the human sel-10  
 CC polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)  
 CC or PS-2, and therefore may be useful for the prevention or treatment of  
 CC Alzheimer's disease.  
 XX  
 SQ Sequence 545 AA;

Query Match 91.7%; Score 540; DB 20; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 MRRKLDHGSEVRSFSLGKKPKCKVSEYTSITGLVPCSATPTTFGDLRAANGQGORRRITS 109  
 DB 6 mrrkldhgsevrslgkpkckvseysttglvpcsatpttfgdlraangqgqrrrits 65  
 QY 110 VQPPTGLQWLKMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQFQDFISLLPKE 169  
 DB 66 vqpptglqewlkmfqsqgpekllalidelidsceptqvkhhmqvleqpqfdrfislilpke 125  
 QY 170 LALYVLSFLEPKDLLQAATCRYWRILAEEDNLLWRECKEEDIDELPHIKRRKVIKPGFI 229  
 DB 126 lalyvlsflepkdllqaactcrywriilaednllwrekekeegideplhikrrkvikpgfi 185  
 QY 230 HSPWKSAIYRQHRIDNWRGELKSPKVLKGDHDDHVTITCLQFCGNRIVSGSDNTLKVWS 289  
 DB 186 hspwksaiyrqhridnwrngelkspkvlkgdhddhvtitclqfcgnrivsgsdntlkvws 245  
 QY 290 AVTGKCLRTLVTGHTGCVWSSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRGM 349  
 DB 246 avtgkclrtlvtghtgcvwssqmrndniisgstdrtlkvwnaetgecihtlyghtstvrGM 305  
 QY 350 HLHEKRVVSGSRDATLRVWMDIETGCQLHVLGMHVAARVCQYDGRVRSVSGAYDFMVKVWD 409  
 DB 306 hlhekrvvsgsrdatlrvmwdietgqclhvlgmhvaarvcydgrrrvsvsgaydfmfvkvwD 365  
 QY 410 PETETCLHTLQHTNRVYSLQFDGIHVVSGLDTSIRVWMDVETGNCIHTLTGHQSLTSGM 469  
 DB 366 petetclhtlqghtnrvyslqfdgihvvsldtsirvwdvetgncihtltghqsltsGM 425  
 QY 470 ELKDNILVSGNADSTVKIWDIKTGQCQLQTLQGNPKHQSATVCLQFNKNFVITSSDDGTVK 529  
 DB 426 elkdnilvsgnadstvkiwdiktgqclqltqgnpkhqsavtclqfnknfvitssddgtvk 485  
 QY 530 LWDLKTGEFIRNLVTLES GSGGVVWRIRASNTKLVCVSGSRNGTEETKLLVLDLDFVDMK 589  
 DB 486 lwdlktgefirnlvtlesgsggvvwrirasnktlvcavsgsrngteetkllvldfdvdmk 545

RESULT 11  
 AAB59196  
 ID AAB59196 standard; protein; 545 AA.  
 XX  
 AC AAB59196;

XX 23-MAR-2001 (first entry)  
XX Human hippocampal sel-10-4 protein.  
XX Sel-10; human; Alzheimer's disease; Abeta.  
XX Homo sapiens.  
XX WO200075328-A1.  
XX 14-DEC-2000.  
XX 23-MAY-2000; 2000WO-US09814.  
XX 09-JUN-1999; 99US-0328877.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Pauley AM, Gurney ME, Li J;  
XX WPI; 2001-102404/11.  
XX New human sel-10 polypeptides and their encoding polynucleotides,  
XX useful for raising antibodies for detecting sel-10 polypeptide  
XX expression and as drug targets in the treatment of Alzheimer's disease  
XX  
XX Claim 1; Page 72-75; 116pp; English.  
XX  
XX The present invention relates to human sel-10. The sel-10 proteins of  
XX the invention are useful for raising monoclonal or polyclonal  
XX antibodies useful in diagnostic assays for detecting sel-10  
XX polypeptide expression. The sel-10 polypeptides are also useful as drug  
XX targets for decreasing antibody levels in the treatment of Alzheimer's  
XX disease. It is also useful for identifying agents capable of  
XX altering the production level of Abeta. The polynucleotides are useful  
XX for developing assays for identifying agents capable of interfering  
XX with the biological pathways that lead to Alzheimer's disease.  
XX  
XX Sequence 545 AA;  
XX  
XX Query Match 91.7%; Score 540; DB 22; Length 545;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 50 MKRKLHGSEVRSLGKKPKCKVSEYTTTGLVPCSATPTTFGDLRAANGOGQORRRITS 109  
XX 6 mkrklhgsevrslgkpkckvseytsttglvpcsatpttfgdlraangogqorrrits 65  
XX  
XX 110 VOPPTGLQEWLKMFSQSGPSEKLLALDELIDSCPTQVKHMQVIEPQFQDFISLLPKE 169  
XX 66 vqpptglqewlkmfsgsgpekllaldelidsceptqvkhumqvlepfqdrdfisllpke 125  
XX  
XX 170 LALYVLSFLEPKDLQAQAOTCRYWRILAEADNLLWRECKEEDPLHKKRKVKPGFI 229  
XX 126 lalyvlsflepkdlqaqaotcrywrilaednllwreckeegideplhkrkvikpgfi 185  
XX  
XX 230 HSPWKSAYIROHRTDNNRRGELSPKVLKGDDHVVITCLOFCGNRIVSGSDNTLKVWS 289  
XX 186 hspwksayirhrtidnnrrgelspkvlkgddhvvitclocfgnrivsgsdntlkvws 245  
XX  
XX 290 AVTGKCLRTLVGHTGGVWSSQMRDNIISGSTDRTLTKVWNAETGECIHTLYGHTSTVRGM 349  
XX 246 avtgkclrtlvghtggvswsqmrdniisgstdrtltkwnaetgecihtlyghtstvr 305  
XX  
XX 350 HLHEKRVVSGSRDFTLRFVMDIETGCLHVLGMHVAARVCQYDGRVVRVSGAYDFMVKVMD 409  
XX 306 hlhekrvvsgsrdftrlrfvmdietgclhvlgmhvaarvcvydgrvvrvsgaydfm 365  
XX  
XX 410 PETETCLHTLQGHNRVYSLQFDGIHVVSGLDTSIRVWDETGNCTIHTLGHQSLSGSM 469  
XX  
XX

Db 366 petetclhtlqghnrvyslqfdgihvvsgsldtsirvwdvetgncihtltghqsltsgm 425  
Qy 470 ELKDNILVSGNADSTVKIWDIKTGQCLOTLOGPNKHQSAVTCLOFNKNFVITSSDDGTVK 529  
Db 426 elkdnilvsgnadstvkIwdiktgqclqlggnpkhgsavtclqfnknfvitssddgtvk 485  
Qy 530 LWDLKTGEFIRNLVTLESQSGGVVWRIRASNTKLVCAGSRNGTEETKLLVLDFDVMK 589  
Db 486 lwdlktgefirnlvtlesqsggvvwrirasantklvcavsrngteetkllvldfdvdmk 545  
RESULT 12  
AAY22463  
ID AAY22463 standard; Protein: 553 AA.  
XX AC AAY22463;  
XX 29-SEP-1999 (first entry)  
XX Human hippocampal sel-10 protein sequence.  
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;  
XX therapy.  
XX Homo sapiens.  
XX WO9932623-A1.  
XX 01-JUL-1999.  
XX 17-DEC-1998; 98WO-US26820.  
XX 19-DEC-1997; 97US-0068243.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Gurney ME, Li J, Pauley AM;  
XX WPI; 1999-458026/38.  
XX N-PSDB; AAX99701.  
XX New isolated human sel-10 polypeptides  
XX Claim 24; Page 50-53; 91pp; English.  
XX  
XX This sequence represents a human sel-10 protein of the invention. This  
XX sequence is specifically a human hippocampal sel-10 protein. The  
XX polypeptides can be used to alter presenilin function. Compounds which  
XX inhibit either the expression or the activity of the human sel-10  
XX polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)  
XX or PS-2, and therefore may be useful for the prevention or treatment of  
XX Alzheimer's disease.  
XX  
XX Sequence 553 AA;  
XX  
XX Query Match 91.7%; Score 540; DB 20; Length 553;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 50 MKRKLHGSEVRSLGKKPKCKVSEYTTTGLVPCSATPTTFGDLRAANGOGQORRRITS 109  
XX 14 mkrklhgsevrslgkpkckvseytsttglvpcsatpttfgdlraangogqorrrits 73  
XX  
XX 110 VOPPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKHMQVIEPQFQDFISLLPKE 169  
XX 74 vqpptglqewlkmfsgsgpekllaldelidsceptqvkhumqvlepfqdrdfisllpke 133  
XX  
XX 170 LALYVLSFLEPKDLQAQAOTCRYWRILAEADNLLWRECKEEDPLHKKRKVKPGFI 229  
XX 134 lalyvlsflepkdlqaqaotcrywrilaednllwreckeegideplhkrkvikpgfi 193  
XX  
XX 230 HSPWKSAYIROHRTDNNRRGELSPKVLKGDDHVVITCLOFCGNRIVSGSDNTLKVWS 289  
XX

|||||  
Db 194 hspwksayirhrdtnwrrgelkspkvlghddhvitclqfcgnrivsgsddntlkvws 253  
QY 290 AVTGKCLRLTGLVHTGCGVWSQMRDNIISGSDRTLKVNNAETGECIHTLYGHTSTVRCM 349  
Db 254 avtgkclrltvghtgvgwssqmrndniisgstdrtlkvwnaetgecihtlyghtstvrcom 313  
QY 350 HLHEKRVSGSRDATLRVWDIETGQCLHVLGMHVAARVCQYDGRVRSVSGAYDFMVKVWD 409  
Db 314 hlhekrvsgsrdatlrvwdietgqclhvlgmhvaarvcydgrrvrsvsgaydfmvmkvwd 373  
QY 410 PETETCLHTLOGHTNRVYSLOFDGHHVWSGLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469  
Db 374 petetclhlqghtnrsvslqfdghvwsidsrsvwdvetgncihtltghqsltsm 433  
QY 470 ELKDNILVSGNADSTVKIWDITGQCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVK 529  
Db 434 elkdnilvsgnadstvkldwtgqclqltqgnpkhqsvatclqfnknfvitssddgtvk 493  
QY 530 LMDLKTGEFIRNLVLTLESQSGGVVWRIRASNTKLYCAVSRNGTEETKLLVLDFFVDMK 589  
Db 494 lwdlktgefirnlvltlesqsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 553

RESULT 13  
AAB59195  
ID AAB59195 standard; protein; 553 AA.  
AC AAB59195;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Human hippocampal sel-10-3 protein.  
XX  
KW Sel-10; human; Alzheimer's disease; Abeta.

OS Homo sapiens.  
XX  
PN WO200075328-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 23-MAY-2000; 2000WO-US09814.  
XX  
PR 09-JUN-1999; 99US-0328877.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Pauley AM, Gurney ME, Li J;  
XX  
DR WPI; 2001-102404/11.  
XX  
PT New human sel-10 polypeptides and their encoding polynucleotides,  
PT useful for raising antibodies for detecting sel-10 polypeptide  
PT expression and as drug targets in the treatment of Alzheimer's disease  
XX  
PS Claim 1; Page 68-71; 116pp; English.  
XX  
CC The present invention relates to human sel-10. The sel-10 proteins of  
CC the invention are useful for raising monoclonal or polyclonal  
CC antibodies useful in diagnostic assays for detecting sel-10  
CC polypeptide expression. The sel-10 polypeptides are also useful as drug  
CC targets for decreasing antibody levels in the treatment of Alzheimer's  
CC disease. It is also useful for identifying agents capable of  
CC altering the production level of Abeta. The polynucleotides are useful  
CC for developing assays for identifying agents capable of interfering  
CC with the biological pathways that lead to Alzheimer's disease.  
XX  
SQ Sequence 553 AA;

Query Match 91.7%; Score 540; DB 22; Length 553;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 50 MKRKLHDGSEVRSFSLGKPKCKVSEYTSITGLVPCSAATPTTFGDLRAANGCQORRITS 109  
Db 14 mkrklhdgsevrfsfslgkpkckvseytsitglvpcsatpttfgdldraangcqqrrits 73  
QY 110 VOPPTGLQEWLKMFSQSWSGPEKLLALDELIDSCPTQVKHMMOVIEPOFORDFISLLPKE 169  
Db 74 vopptglqewlkmfswsgpekllalidelidscepqvkhhmmqvlepfqfdrfislpke 133  
QY 170 LALYVLSFLEPKDLQAQTCRYWRILAEENLWRECKEEDIDELPHIKRRKVIKPGFI 229  
Db 134 lalyvlsflepkdlqaactcrywriilaednlwreckeegideplhikrrkvikpgfi 193  
QY 230 HSPWKSAYIRQHRIDTNWRRGELKSPKVLGHDDHVTCLQFCGNRIVSGSDNTLKVWS 289  
Db 194 hspwksayirhrdtnwrrgelkspkvlghddhvitclqfcgnrivsgsddntlkvws 253  
QY 290 AVTGKCLRLTGLVHTGCGVWSQMRDNIISGSDRTLKVNNAETGECIHTLYGHTSTVRCM 349  
Db 254 avtgkclrltvghtgvgwssqmrndniisgstdrtlkvwnaetgecihtlyghtstvrcom 313  
QY 350 HLHEKRVSGSRDATLRVWDIETGQCLHVLGMHVAARVCQYDGRVRSVSGAYDFMVKVWD 409  
Db 314 hlhekrvsgsrdatlrvwdietgqclhvlgmhvaarvcydgrrvrsvsgaydfmvmkvwd 373  
QY 410 PETETCLHTLOGHTNRVYSLOFDGHHVWSGLDTSIRVWDVETGNCIHTLTGHQSLTSGM 469  
Db 374 petetclhlqghtnrsvslqfdghvwsidsrsvwdvetgncihtltghqsltsm 433  
QY 470 ELKDNILVSGNADSTVKIWDITGQCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVK 529  
Db 434 elkdnilvsgnadstvkldwtgqclqltqgnpkhqsvatclqfnknfvitssddgtvk 493  
QY 530 LMDLKTGEFIRNLVLTLESQSGGVVWRIRASNTKLYCAVSRNGTEETKLLVLDFFVDMK 589  
Db 494 lwdlktgefirnlvltlesqsggvvwrirasntklvcavgsrngteetkllvldfdvdmk 553

RESULT 14  
AAY22462  
ID AAY22462 standard; Protein; 592 AA.  
XX  
AC AAY22462;  
XX  
DT 29-SEP-1999 (first entry)  
XX  
DE Human hippocampal sel-10 protein sequence.  
XX  
KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;  
KW therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO9932623-A1.  
XX  
PD 01-JUL-1999.  
XX  
PF 17-DEC-1998; 98WO-US26820.  
XX  
PR 19-DEC-1997; 97US-0068243.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Gurney ME, Li J, Pauley AM;  
XX  
DR WPI; 1999-458026/38.  
XX  
DR N-PSDB; AAX9701.  
XX  
PT New isolated human sel-10 polypeptides  
XX  
PS Claim 24; Page 47-50; 91pp; English.

XX This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human hippocampal sel-10 protein. The  
CC polypeptides can be used to alter presenilin function. Compounds which  
CC inhibit either the expression or the activity of the human sel-10  
CC polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)  
CC or PS-2, and therefore may be useful for the prevention or treatment of  
CC Alzheimer's disease.  
XX  
SQ Sequence 592 AA;

Query Match 91.7%; Score 540; DB 20; Length 592;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 MKRLDHGSEVRSFSLGKKPKCVSEYTTGLVPCSAPTTFGLRAANGOGQRRRTS 109  
Db 53 mkrldhgsevrsfslgkpkcvseysttglvpcsatpttfgdlraangggqrrrits 112  
Qy 110 VOPPTGLQEWLKMFSQSGPKLLALDELIDSCPTQVKHMMQVIEPQFQDFISLLPKE 169  
Db 113 vppptglqewlkmfsgsgpekllaldelidsceptqvkmmqvlepfqfdrfslpke 172  
Qy 170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEIGIDEPGLHKRRKVIKPGFI 229  
Db 173 lalyvlsflepkdllqaatcrywriilaednllwrekckeegideplhkrkrvikpgfi 232  
Qy 230 HSPWKSAYIROHRIDTNWRRGELSPKVLKGDDHVTCLQFCGNRIYSGSDNTLKWS 289  
Db 233 hspwksayirhridtnwrrgelspkvlkgddhvtclqfcgnriyvsddntlkws 292  
Qy 290 AVTGKCLRTLQHTGVVSSQMRDNIISGSTDRTLKVNNAETGECIHTLXGHTSTVRGM 349  
Db 293 avtgkclrtlvghgtgvvssqmrndniisgstdrtlkvwnaetgecihtlyghtstvr 352  
Qy 350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVCVQDGRVRSVGYDFMVKVWD 409  
Db 353 hlhekrvvsgsrdatlrwvdietsgclhvlgmhvaavrcvqdgrrvrsyaydfmrvkwd 412  
Qy 410 PETETCLHTLQHTNRVYSLQFDGIHVHVGSLDTSIRWVDYETGNCIHTLGHQSLTSGM 469  
Db 413 petetclhtlqhtnrvyqlfdgihvvsghltsirwvdyetgncihtlghqsltsgm 472  
Qy 470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPKNKHQSAVTCLOFNKVFVITSSDDGTVK 529  
Db 473 elkdnilvsgnadstvkikwdiktgcclqtlgppnkhqgsavtclqfknfvitssddgtvk 532  
Qy 530 LWDLKTGEFIRNLVTLESQSGGVVWRIRASNTKLVCAGSRNGTEETKLLVDFDVKM 589  
Db 533 lwdlktgefirnlvtlesqsggvvwrirasntklvcavsrngteetkllvldfdvdkm 592

RESULT 15  
AAB59194  
ID AAB59194 standard; protein; 592 AA.  
XX  
AC AAB59194;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
XX Human hippocampal sel-10-2 protein.  
DE  
XX Sel-10; human; Alzheimer's disease; Abeta.  
XX  
XX Homo sapiens.  
OS  
XX WO200075328-A1.  
PN  
XX 14-DEC-2000.  
PD  
XX 23-MAY-2000; 2000WO-US09814.  
PF  
XX

PR 09-JUN-1999; 99US-0328877.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
PA  
XX Pauley AM, Gurney ME, Li J;  
PI  
XX WPI; 2001-102404/11.  
DR  
XX New human sel-10 polypeptides and their encoding polynucleotides,  
PT useful for raising antibodies for detecting sel-10 polypeptide  
PT expression and as drug targets in the treatment of Alzheimer's disease  
PT  
XX  
PS Claim 1; Page 64-68; 116pp; English.  
XX  
CC The present invention relates to human sel-10. The sel-10 proteins of  
CC the invention are useful for raising monoclonal or polyclonal  
CC antibodies useful in diagnostic assays for detecting sel-10  
CC polypeptide expression. The sel-10 polypeptides are also useful as drug  
CC targets for decreasing antibody levels in the treatment of Alzheimer's  
CC disease. It is also useful for identifying agents capable of  
CC altering the production level of Abeta. The polynucleotides are useful  
CC for developing assays for identifying agents capable of interfering  
CC with the biological pathways that lead to Alzheimer's disease.  
XX  
SQ Sequence 592 AA;

Query Match 91.7%; Score 540; DB 22; Length 592;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 MKRLDHGSEVRSFSLGKKPKCVSEYTTGLVPCSAPTTFGLRAANGOGQRRRTS 109  
Db 53 mkrldhgsevrsfslgkpkcvseysttglvpcsatpttfgdlraangggqrrrits 112  
Qy 110 VOPPTGLQEWLKMFSQSGPKLLALDELIDSCPTQVKHMMQVIEPQFQDFISLLPKE 169  
Db 113 vppptglqewlkmfsgsgpekllaldelidsceptqvkmmqvlepfqfdrfslpke 172  
Qy 170 LALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWREKCKEIGIDEPGLHKRRKVIKPGFI 229  
Db 173 lalyvlsflepkdllqaatcrywriilaednllwrekckeegideplhkrkrvikpgfi 232  
Qy 230 HSPWKSAYIROHRIDTNWRRGELSPKVLKGDDHVTCLQFCGNRIYSGSDNTLKWS 289  
Db 233 hspwksayirhridtnwrrgelspkvlkgddhvtclqfcgnriyvsddntlkws 292  
Qy 290 AVTGKCLRTLQHTGVVSSQMRDNIISGSTDRTLKVNNAETGECIHTLXGHTSTVRGM 349  
Db 293 avtgkclrtlvghgtgvvssqmrndniisgstdrtlkvwnaetgecihtlyghtstvr 352  
Qy 350 HLHEKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVCVQDGRVRSVGYDFMVKVWD 409  
Db 353 hlhekrvvsgsrdatlrwvdietsgclhvlgmhvaavrcvqdgrrvrsyaydfmrvkwd 412  
Qy 410 PETETCLHTLQHTNRVYSLQFDGIHVHVGSLDTSIRWVDYETGNCIHTLGHQSLTSGM 469  
Db 413 petetclhtlqhtnrvyqlfdgihvvsghltsirwvdyetgncihtlghqsltsgm 472  
Qy 470 ELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPKNKHQSAVTCLOFNKVFVITSSDDGTVK 529  
Db 473 elkdnilvsgnadstvkikwdiktgcclqtlgppnkhqgsavtclqfknfvitssddgtvk 532  
Qy 530 LWDLKTGEFIRNLVTLESQSGGVVWRIRASNTKLVCAGSRNGTEETKLLVDFDVKM 589  
Db 533 lwdlktgefirnlvtlesqsggvvwrirasntklvcavsrngteetkllvldfdvdkm 592

Search completed: September 25, 2001, 14:45:34  
Job time: 209 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:43:40 ; Search time 13.51 Seconds  
(without alignments)  
897.684 Million cell updates/sec

Title: US-09-328-877A-8  
Perfect score: 589  
Sequence: 1 MSKPGKPTLNHGLVPVDLKS.....SRNGTEETKLVLDFDVMK 589

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	2.0	209	3	US-08-899-578-6
2	12	2.0	587	3	US-08-899-578-2
3	8	1.4	13	2	US-08-671-094B-3
4	8	1.4	13	2	US-08-671-094B-5
5	8	1.4	14	2	US-08-671-094B-2
6	8	1.4	15	1	US-08-311-307B-6
7	8	1.4	16	1	US-08-311-307B-3
8	8	1.4	16	1	US-08-311-307B-4
9	8	1.4	37	3	US-08-476-376-12
10	8	1.4	73	1	US-07-792-988-3
11	8	1.4	73	1	US-07-792-988-4
12	8	1.4	73	1	US-07-778-413E-18
13	8	1.4	73	1	US-07-778-413E-19
14	8	1.4	73	1	US-07-778-413E-20
15	8	1.4	73	1	US-08-340-102-18
16	8	1.4	73	1	US-08-340-102-19
17	8	1.4	73	1	US-08-340-102-20
18	8	1.4	73	1	US-08-330-163-2
19	8	1.4	73	1	US-08-330-163-3
20	8	1.4	73	1	US-08-330-163-4
21	8	1.4	73	1	US-08-482-111-2
22	8	1.4	73	1	US-08-482-111-3
23	8	1.4	73	1	US-08-482-111-4
24	8	1.4	73	2	US-08-436-420-26
25	8	1.4	73	2	US-08-436-420-27
26	8	1.4	73	2	US-08-436-420-28
27	8	1.4	73	3	US-08-846-966-2

28	8	1.4	73	3	US-08-846-966-3	Sequence 3, Appli
29	8	1.4	73	3	US-08-846-966-4	Sequence 4, Appli
30	8	1.4	73	3	US-08-557-142-2	Sequence 2, Appli
31	8	1.4	73	3	US-08-557-142-3	Sequence 3, Appli
32	8	1.4	73	3	US-08-557-142-4	Sequence 4, Appli
33	8	1.4	73	5	PCT-US94-06264-2	Sequence 2, Appli
34	8	1.4	73	5	PCT-US94-06264-3	Sequence 3, Appli
35	8	1.4	73	5	PCT-US94-06264-4	Sequence 4, Appli
36	8	1.4	106	1	US-08-352-324A-5	Sequence 5, Appli
37	8	1.4	106	2	US-08-862-607-5	Sequence 4, Appli
38	8	1.4	106	2	US-08-468-819-7	Sequence 7, Appli
39	8	1.4	106	3	US-09-203-235-5	Sequence 7, Appli
40	8	1.4	106	5	PCT-US95-16144-5	Sequence 5, Appli
41	8	1.4	107	1	US-08-352-324A-4	Sequence 4, Appli
42	8	1.4	107	1	US-08-352-324A-7	Sequence 7, Appli
43	8	1.4	107	2	US-08-862-607-4	Sequence 4, Appli
44	8	1.4	107	2	US-08-862-607-7	Sequence 7, Appli
45	8	1.4	107	2	US-08-468-819-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-899-578-6  
; Sequence 6, Application US/08899578  
; Patent No. 6087153  
; GENERAL INFORMATION:  
; APPLICANT: Greenwald, Iva  
; APPLICANT: Hubbard, E. Jane  
; TITLE OF INVENTION: SEL-10 AND USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/899,578  
; FILING DATE: 24-JUL-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/ARC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 278-0525  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 209 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-899-578-6

Query Match 2.0%; Score 12; DB 3; Length 209;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 438 SGLDFTSIRWMD 449  
Db 139 SGLDFTSIRWMD 150

```

RESULT 2
US-08-899-578-2
; Sequence 2, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-899-578-2

Query Match 2.0%; Score 12; DB 3; Length 587;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 438 SGSLDTSIRVWD 449
Db 435 SGSLDTSIRVWD 446
|||||
|||||

RESULT 3
US-08-671-094B-3
; Sequence 3, Application US/08671094B
; Patent No. 5912232
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Anti-inflammatory Polypeptide
; TITLE OF INVENTION: Antagonists of Human Interleukin-8
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid

```

```

; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues
; OTHER INFORMATION: may be substituted with aminobutyric acid, homocysteine, o
; OTHER INFORMATION: diaminosuberic acid."
; US-08-671-094B-3

Query Match 1.4%; Score 8; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 494 QCLQTFLQ 501
Db 3 QCLQTFLQ 10
|||||

RESULT 4
US-08-671-094B-5
; Sequence 5, Application US/08671094B
; Patent No. 5912232
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Anti-inflammatory Polypeptide
; TITLE OF INVENTION: Antagonists of Human Interleukin-8
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid

```

STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 4..6  
OTHER INFORMATION: /note= "The cysteine residues  
may be substituted with aminobutyric acid, homocysteine, or  
diaminosuberic acid."  
US-08-671-094B-5

Query Match 1.4%; Score 8; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.76; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0

QY 494 QCLQTLQG 501  
| | | | |  
DB 3 QCLQTLQG 10

RESULT 5  
US-08-671-094B-2  
; Sequence 2, Application US/08671094B  
; Patent No. 5912232  
; GENERAL INFORMATION:  
; APPLICANT: Talmadge, James E.  
; TITLE OF INVENTION: Anti-Inflammatory Polypeptide  
; TITLE OF INVENTION: Antagonists of Human Interleukin-8  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
; STREET: 1601 Market Street  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2307

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/671,094B  
FILING DATE: 28-JUN-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hagan, Patrick J.  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 63086FWC  
TELEPHONE: (215)563-4100  
TELEFAX: (215)563-4044  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 4..6  
OTHER INFORMATION: /note= "The cysteine residues  
may be substituted with aminobutyric acid, homocysteine, or  
diaminosuberic acid."  
US-08-671-094B-2

Query Match 1.4%; Score 8; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQG 501  
| | | | |  
DB 3 QCLQTLQG 10

RESULT 6  
US-08-311-307B-6  
; Sequence 6, Application US/08311307B  
; Patent No. 5627156  
; GENERAL INFORMATION:  
; APPLICANT: Talmadge, James E.  
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human  
; TITLE OF INVENTION: Interleukin-8  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
; STREET: 1601 Market Street  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2307

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,307B  
FILING DATE: 23-SEP-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hagan, Patrick J.  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 63085  
TELEPHONE: (215)563-4100  
TELEFAX: (215)563-4044  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 4..6  
OTHER INFORMATION: /note= "The cysteine residues may  
be substituted with aminobutyric acid, homocysteine, or  
diaminosuberic acid."  
US-08-311-307B-6

Query Match 1.4%; Score 8; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQG 501  
| | | | |  
DB 5 QCLQTLQG 12

RESULT 7  
US-08-311-307B-3  
; Sequence 3, Application US/08311307B  
; Patent No. 5627156  
; GENERAL INFORMATION:

```

; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; TITLE OF INVENTION: Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; OTHER INFORMATION: be substituted with aminobutyric acid, homocysteine, or
; OTHER INFORMATION: diaminosuberic acid."
;
US-08-311-307B-3

```

```

Query Match      1.4%; Score 8; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 494 QCLQTLOG 501
      |||||
Db 5 QCLQTLOG 12

```

```

RESULT 8
US-08-311-307B-4
; Sequence 4, Application US/08311307B
; Patent No. 5627156
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; TITLE OF INVENTION: Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; OTHER INFORMATION: be substituted with aminobutyric acid, homocysteine or
; OTHER INFORMATION: diaminosuberic acid."
;
US-08-311-307B-4

```

```

Query Match      1.4%; Score 8; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 494 QCLQTLOG 501
      |||||
Db 5 QCLQTLOG 12

```

```

RESULT 9
US-08-476-376-12
; Sequence 12, Application US/08476376
; Patent No. 6103234
; GENERAL INFORMATION:
; APPLICANT: WOLPE, STEPHEN D.
; APPLICANT: CERAMI, ANTHONY
; APPLICANT: SHERRY, BARBARA
; APPLICANT: TEKAMP-OLSON, PATRICIA A.
; TITLE OF INVENTION: MACROPHAGE-DERIVED INFLAMMATORY MEDIATOR
; TITLE OF INVENTION: (MIP-2)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID A. JACKSON, ESQ.
; STREET: 411 HACKENSACK AVE, CONTINENTAL PLAZA, 4TH
; STREET: FLOOR
; CITY: HACKENSACK
; STATE: NEW JERSEY
; COUNTRY: U.S.A.
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,376
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,498
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/105,105  
FILING DATE: 10-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/914,045  
FILING DATE: 13-JUL-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/399,971  
FILING DATE: 01-SEP-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/240,078  
FILING DATE: 02-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/104,827  
FILING DATE: 02-OCT-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/766,852  
FILING DATE: 16-AUG-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/414,098  
FILING DATE: 07-SEP-1982  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/351,290  
FILING DATE: 22-FEB-1982  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/299,932  
FILING DATE: 08-SEP-1981  
ATTORNEY/AGENT INFORMATION:  
NAME: JACKSON ESQ., DAVID A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 13046 CIP A2D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-476-376-12

Query Match 1.4%; Score 8; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQG 501  
Db 15 QCLQTLQG 22

RESULT 10  
US-07-792-988-3  
Sequence 3, Application US/07792988  
Patent No. 5306709  
GENERAL INFORMATION:  
APPLICANT: Gewirtz, Alan M.  
TITLE OF INVENTION: Suppression of megakaryo-  
cytopenia by macrophage inflammatory proteins  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: University of Pennsylvania  
STREET: Suite 419  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/792,988  
FILING DATE: 19911115  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-159  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5306709e  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
US-07-792-988-3

Query Match 1.4%; Score 8; DB 1; Length 73;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQG 501  
Db 10 QCLQTLQG 17

RESULT 11  
US-07-792-988-4  
Sequence 4, Application US/07792988  
Patent No. 5306709  
GENERAL INFORMATION:  
APPLICANT: Gewirtz, Alan M.  
TITLE OF INVENTION: Suppression of megakaryo-  
cytopenia by macrophage inflammatory proteins  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: University of Pennsylvania  
STREET: Suite 419  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/792,988  
FILING DATE: 19911115  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-159  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5306709e

```
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
US-07-792-988-4

Query Match 1.4%; Score 8; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 494 QCLQTLOG 501
Db 10 QCLQTLOG 17

RESULT 12
US-07-778-413E-18
; Sequence 18, Application US/07778413E
; Patent No. 5401651
; GENERAL INFORMATION:
; APPLICANT: Walz, Alfred
; TITLE OF INVENTION: No. 5401651el Neutrophil
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 2.0 MB
; OPERATING SYSTEM: Macintosh OS 7.0
; SOFTWARE: Microsoft Word Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07778,413E
; FILING DATE: 16-OCT-1991
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook, Robert R.
; REGISTRATION NUMBER: 31602
; REFERENCE/DOCKET NUMBER: A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (805) 499-5725 EXTENSION 4955
; TELEFAX: (805) 499-8011
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-778-413E-19

Query Match 1.4%; Score 8; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 494 QCLQTLOG 501
Db 10 QCLQTLOG 17

RESULT 14
US-07-778-413E-20
; Sequence 20, Application US/07778413E
; Patent No. 5401651
; GENERAL INFORMATION:
; APPLICANT: Walz, Alfred
; TITLE OF INVENTION: No. 5401651el Neutrophil
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 2.0 MB
; OPERATING SYSTEM: Macintosh OS 7.0
; SOFTWARE: Microsoft Word Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07778,413E
; FILING DATE: 16-OCT-1991
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook, Robert R.
; REGISTRATION NUMBER: 31602
; REFERENCE/DOCKET NUMBER: A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (805) 499-5725 EXTENSION 4955
; TELEFAX: (805) 499-8011
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-778-413E-18

Query Match 1.4%; Score 8; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 494 QCLQTLOG 501
Db 10 QCLQTLOG 17

RESULT 13
US-07-778-413E-19
; Sequence 19, Application US/07778413E
; Patent No. 5401651
```

ATTORNEY/AGENT INFORMATION:  
NAME: Cook, Robert R.  
REGISTRATION NUMBER: 31602  
REFERENCE/DOCKET NUMBER: A-204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (805) 499-5725 EXTENSION 4955  
TELEFAX: (805) 499-8011  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-778-413E-20

Query Match 1.4%; Score 8; DB 1; Length 73;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 OCLQTLQG 501

Db 10 OCLQTLQG 17

## RESULT 15

US-08-340-102-18  
Sequence 18, Application US/08340102  
Patent No. 5591718  
GENERAL INFORMATION:  
APPLICANT: Walz, Alfred  
TITLE OF INVENTION: No. 5591718el Neutrophil  
TITLE OF INVENTION: Activating Factors  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: Amgen Center  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB  
COMPUTER: MS-DOS  
OPERATING SYSTEM: MS-DOS 6.22  
SOFTWARE: Microsoft Word Version 5.1a for  
SOFTWARE: Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,102  
FILING DATE: 15-NOV-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Cook, Robert R.  
REGISTRATION NUMBER: 31602  
REFERENCE/DOCKET NUMBER: A-204A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (805) 499-5725 EXTENSION 4955  
TELEFAX: (805) 499-8011  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-340-102-18

Query Match 1.4%; Score 8; DB 1; Length 73;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 OCLQTLQG 501

Db 10 OCLQTLQG 17

Search completed: September 25, 2001, 14:45:58  
Job time: 138 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:44:10 ; Search time 17.49 seconds  
(without alignments)  
2565.284 Million cell updates/sec

Title: US-09-328-877A-8

Perfect score: 589

Sequence: 1 MSKPGKPTLNHGLVPVLDLKS.....SRNGTEETKLLVDFDVMK 589

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

PIR68:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	2.0	579	2 T22703	hypothetical prote
2	9	1.5	524	2 T46310	hypothetical prote
3	8	1.4	33	2 S50035	cytokine bogro - b
4	8	1.4	107	2 B38290	GRO-gamma precurs
5	8	1.4	107	2 A28414	melanoma growth-st
6	8	1.4	107	2 JH0281	macrophage inflamm
7	8	1.4	272	2 H70335	hypothetical prote
8	8	1.4	330	2 S27707	daunorubicin resis
9	8	1.4	336	2 T27762	hypothetical prote
10	8	1.4	344	2 I45774	odorant receptor 1
11	8	1.4	386	2 T26805	hypothetical prote
12	8	1.4	515	2 A43289	CDC4L protein - hu
13	8	1.4	701	2 T16607	hypothetical prote
14	8	1.4	821	2 A86243	hypothetical prote
15	7	1.2	40	2 S56768	capsid protein - L
16	7	1.2	107	2 JQ2035	hypothetical 11.7K
17	7	1.2	123	2 S29121	GTP-binding regula
18	7	1.2	135	2 T03288	ferredoxin [2Fe-2S
19	7	1.2	149	2 S33208	serine proteinase
20	7	1.2	149	2 T23179	hypothetical prote
21	7	1.2	164	2 T33915	hypothetical prote
22	7	1.2	187	2 A72737	hypothetical prote
23	7	1.2	190	2 T35570	hypothetical prote
24	7	1.2	196	2 T26943	hypothetical prote
25	7	1.2	206	2 H81932	probable dTMP kin
26	7	1.2	206	2 B81173	thymidylate kinase
27	7	1.2	209	2 G71430	hypothetical prote
28	7	1.2	221	2 T10221	hypothetical prote
29	7	1.2	227	2 T21783	hypothetical prote

30	7	1.2	232	2 T21359	hypothetical prote
31	7	1.2	249	2 A96632	hypothetical prote
32	7	1.2	251	2 B70450	beta 1,4 glucosylt
33	7	1.2	253	2 T16918	hypothetical prote
34	7	1.2	255	2 S55641	uracil DNA glycosy
35	7	1.2	268	2 S76063	hypothetical prote
36	7	1.2	273	2 H86241	protein T16B5.4 [i
37	7	1.2	273	2 S61637	probable membrane
38	7	1.2	289	2 H82364	O-methyltransferas
39	7	1.2	298	2 C83216	probable epoxide h
40	7	1.2	304	2 T40316	conserved kin7-lik
41	7	1.2	310	2 T43158	probable GTP-bindi
42	7	1.2	314	2 T43299	probable GTP-bindi
43	7	1.2	319	2 A70340	conserved hypotet
44	7	1.2	319	2 A72637	hypothetical prote
45	7	1.2	326	1 RGOB02	GTP-binding regula

#### ALIGNMENTS

RESULT 1

T22703

hypothetical protein F55B12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22703

R:Slms, M.

submitted to the EMBL Data Library, September 1996

A:Reference number: Z19602

A:Accession: T22703

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-579 <WIL>

A:Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GN00023; CESP:F55B12.3

A:Experimental source: clone F55B12

C:Genetics:

A:Gene: CESP:F55B12.3

A:Map position: 5

A:Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 2.0%; Score 12; DB 2; Length 579;  
Best Local Similarity 100.0%; Pred. No. 0.00071;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 SGSLDTSIRVWD 449

|||||

Db 427 SGSLDTSIRVWD 438

RESULT 2

T46310

hypothetical protein DKF2p434G0511.1 - human

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

C:Accession: T46310

R:Duesterhoft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23035

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-524 <AAA>

A:Cross-references: EMBL:AL137748

A:Experimental source: adult testis; clone DKF2p434G0511

C:Genetics:

A:Note: DKF2p434G0511.1

Query Match

Best Local Similarity 1.5%; Score 9; DB 2; Length 524;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 VSGSRDATL 365  
 |||||  
 Db 318 VSGSRDATL 326

RESULT 3  
 S50035  
 cytokine boGRO - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 13-Jan-1995 #sequence\_revision 12-May-1995 #text\_change 02-Jun-1995  
 C:Accession: S50035  
 R:Rogivue, C.; Car, B.D.; Allmann-Iselin, I.; Zwahlen, R.D.; Walz, A.  
 submitted to the Protein Sequence Database, December 1994  
 A:Description: Bovine GRO (boGRO), a new monocyte-macrophage-derived cytokine of the int  
 A:Reference number: S50035  
 A:Accession: S50035  
 A:Molecule type: protein  
 A:Residues: 1-33 <ROG>  
 C:Superfamily: beta-thromboglobulin

Query Match 1.4%; Score 8; DB 2; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.88;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501  
 |||||  
 Db 10 QCLQTLOG 17

RESULT 4  
 B38290  
 GRO-gamma precursor - human  
 N:Alternate names: growth-regulated protein gamma; macrophage inflammatory protein 2 bet  
 C:Species: Homo sapiens (man)  
 C:Date: 31-May-1991 #sequence\_revision 27-Oct-1995 #text\_change 20-Aug-1999  
 C:Accession: JH0282; B38290; C46519  
 R:Tekamp-Olson, P.; Gallegos, C.; Bauer, D.; McClain, J.; Sherry, B.; Fabre, M.; van Dev  
 J. Exp. Med. 172, 911-919, 1990  
 A:Title: Cloning and characterization of cDNAs for murine macrophage inflammatory protei  
 A:Reference number: JH0200; MUID:90354792  
 A:Accession: JH0282  
 A:Molecule type: mRNA  
 A:Residues: 1-107 <TEK>  
 A:Cross-references: GB:X53800; NID:g34662; PIDN:CAA37809.1; PID:g34663  
 R:Haskill, S.; Peace, A.; Morris, J.; Sporn, S.A.; Anisowicz, A.; Lee, S.W.; Smith, T.;  
 Proc. Natl. Acad. Sci. U.S.A. 87, 7732-7736, 1990  
 A:Title: Identification of three related human GRO genes encoding cytokine functions.  
 A:Reference number: A38290; MUID:91017578  
 A:Accession: B38290  
 A:Molecule type: mRNA  
 A:Residues: 1-26,'G',29-107 <HAS>  
 A:Cross-references: GB:M36821; NID:g183632; PIDN:AAA63184.1; PID:g183633  
 R:Proost, P.; De Wolf-Peters, C.; Conings, R.; Opdenakker, G.; Billiau, A.; Van Damme,  
 J. Immunol. 150, 1000-1010, 1993  
 A:Title: Identification of a novel granulocyte chemotactic protein (GCP-2) from human tu  
 A:Reference number: A46519; MUID:93139489  
 A:Accession: C46519  
 A:Molecule type: protein  
 A:Residues: 35-52 <PRO>  
 A:Experimental source: MG-63 osteosarcoma cells  
 C:Genetics:  
 A:Map position: 4q21  
 C:Superfamily: beta-thromboglobulin  
 F:1-34/Domain: signal sequence #status predicted <SIG>  
 F:35-107/Product: GRO-gamma #status experimental <MAT>

Query Match 1.4%; Score 8; DB 2; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501  
 |||||  
 Db 44 QCLQTLOG 51

RESULT 5  
 A28414  
 melanoma growth-stimulatory activity precursor - human  
 N:Alternate names: fibroblast-derived neutrophil-activating protein gamma; GRO-alpha;  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 20-Aug-1999  
 C:Accession: SL3669; A28414; S00983; B60401; S03976; A47626; B46519  
 R:Baker, N.E.; Kucera, G.; Richmond, A.  
 Nucleic Acids Res. 18, 6453, 1990  
 A:Title: Nucleotide sequence of the human melanoma growth stimulatory activity (MGSA)  
 A:Reference number: SL3669; MUID:91057157  
 A:Accession: SL3669  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-107 <BAK>  
 A:Cross-references: EMBL:X54489; NID:g34625; PIDN:CAA38361.1; PID:g34626  
 R:Anisowicz, A.; Bardwell, L.; Sager, R.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 7188-7192, 1987  
 A:Title: Constitutive overexpression of a growth-regulated gene in transformed Chinese  
 A:Reference number: A94184; MUID:88041072  
 A:Accession: A28414  
 A:Molecule type: mRNA  
 A:Residues: 1-107 <ANI>  
 A:Cross-references: GB:J03561; NID:g183622; PIDN:AAA35933.1; PID:g306806  
 R:Richmond, A.; Balentien, E.; Thomas, H.G.; Flagg, G.; Barton, D.E.; Spiess, J.; Bo  
 EMBO J. 7, 2025-2033, 1988  
 A:Title: Molecular characterization and chromosomal mapping of melanoma growth stimu  
 A:Reference number: S00983; MUID:88328991  
 A:Accession: S00983  
 A:Molecule type: mRNA  
 A:Residues: 1-107 <RTC>  
 A:Cross-references: EMBL:X12510; NID:g34621; PIDN:CAA31027.1; PID:g34622  
 R:Schoeder, J.M.; Sticherling, M.; Henneicke, H.H.; Preissner, W.C.; Christophers, E  
 J. Immunol. 144, 2223-2232, 1990  
 A:Title: IL-1alpha or tumor necrosis factor-alpha stimulate release of three NAP-1/IL  
 A:Reference number: A60401; MUID:90187866  
 A:Accession: B60401  
 A:Molecule type: protein  
 A:Residues: 35-42,'X',44,'X',46-48 <SCH>  
 A:Experimental source: dermal fibroblasts  
 R:Golds, E.E.; Mason, P.; Nyirkos, P.  
 Biochem. J. 259, 585-588, 1989  
 A:Title: Inflammatory cytokines induce synthesis and secretion of gro protein and a n  
 A:Reference number: S03975; MUID:89246368  
 A:Accession: S03976  
 A:Molecule type: protein  
 A:Residues: 35-41,'X',43-49,'X',51-52,'XX',55-57 <GOL>  
 R:Schoeder, J.M.; Persoon, N.L.M.; Christophers, E.  
 J. Exp. Med. 171, 1091-1100, 1990  
 A:Title: Lipopolysaccharide-stimulated human monocytes secrete, apart from neutrophil  
 ntity with melanoma growth-stimulatory activity.  
 A:Reference number: A47626; MUID:90217938  
 A:Accession: A47626  
 A:Molecule type: protein  
 A:Residues: 35-63,'X',65 <SC2>  
 A:Experimental source: LPS-stimulated monocytes  
 R:Proost, P.; De Wolf-Peters, C.; Conings, R.; Opdenakker, G.; Billiau, A.; Van Damme  
 J. Immunol. 150, 1000-1010, 1993  
 A:Title: Identification of a novel granulocyte chemotactic protein (GCP-2) from human  
 A:Reference number: A46519; MUID:93139489  
 A:Accession: B46519  
 A:Molecule type: protein  
 A:Residues: 35-62 <PRO>  
 A:Experimental source: MG-63 osteosarcoma cells  
 C:Genetics:  
 A:Gene: GDB:GR01  
 A:Cross-references: GDB:L20181; OMIM:155730  
 A:Map position: 4q21-4q21

C:Superfamily: beta-thromboglobulin  
 F;1-34/Domain: signal sequence #status predicted <SIG>  
 F;35-107/Product: melanoma growth-stimulatory activity #status experimental <MAT>

Query Match 1.4% Score 8; DB 2; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501  
 |||||  
 Db 44 QCLQTLOG 51

## RESULT 6

JH0281  
 macrophage inflammatory protein 2 alpha precursor - human  
 N;Alternate names: gro-beta; growth regulated protein beta; melanoma growth-stimulatory  
 C;Species: Homo sapiens (man)  
 C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Aug-1999  
 C;Accession: JH0281; A35931; A38290; A60407  
 R;Takamp-Olson, P.; Gallegos, C.; Bauer, D.; McClain, J.; Sherry, B.; Fabre, M.; van Der  
 J. Exp. Med. 172, 911-919, 1990  
 A;Title: Cloning and characterization of cDNAs for murine macrophage inflammatory protei  
 A;Reference number: JH0200; MUID:90354792

A;Accession: JH0281  
 A;Molecule type: mRNA  
 A;Residues: 1-107 <TEK>  
 A;Cross-references: GB:X53799; NID:g34658; PIDN:CAA37808.1; PID:g34659  
 R;lida, N.; Grotendorst, G.R.  
 Mol. Cell. Biol. 10, 5596-5599, 1990  
 A;Title: Cloning and sequencing of a new gro transcript from activated human monocytes:  
 A;Reference number: A35931; MUID:90377259  
 A;Accession: A35931  
 A;Molecule type: mRNA  
 A;Residues: 1-107 <IID>  
 A;Cross-references: GB:M57731; GB:M36964; NID:q183626; PIDN:AAA63182.1; PID:q183627  
 R;Haskill, S.; Peace, A.; Morris, J.; Sporn, S.A.; Anisowicz, A.; Lee, S.W.; Smith, T.;  
 Proc. Natl. Acad. Sci. U.S.A. 87, 7732-7736, 1990  
 A;Title: Identification of three related human GRO genes encoding cytokine functions.  
 A;Reference number: A38290; MUID:91017578

A;Accession: A38290  
 A;Molecule type: mRNA  
 A;Residues: 1-107 <HAS>  
 A;Cross-references: GB:M36820; NID:g183628; PIDN:AAA63183.1; PID:g183629  
 R;Sporn, S.A.; Eterman, D.F.; Johnson, C.E.; Morris, J.; Martin, G.; Ladner, M.; Haskill  
 J. Immunol. 144, 4434-4441, 1990  
 A;Title: Monocyte adherence results in selective induction of novel genes sharing homolo  
 A;Reference number: A60407; MUID:90257367

A;Accession: A60407  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 56-107 <SPO>  
 C;Superfamily: beta-thromboglobulin  
 C;Keywords: Inflammation  
 F;1-34/Domain: signal sequence #status predicted <SIG>  
 F;35-107/Product: macrophage inflammatory protein 2 alpha #status predicted <MAT>

Query Match 1.4% Score 8; DB 2; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501  
 |||||  
 Db 44 QCLQTLOG 51

## RESULT 7

H70335  
 hypothetical protein aq\_397 - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999

C;Accession: H70335  
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
 V. Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A;Reference number: A70300; MUID:98196666  
 A;Accession: H70335  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-272 <AQF>  
 A;Cross-references: GB:AE000688; NID:g2983063; PIDN:AA06680.1; PID:g2983073; GB:AE00  
 A;Experimental source: strain VF5  
 C;Genetics:  
 A;Gene: aq\_397  
 C;Superfamily: Aquifex aeolicus hypothetical protein aq\_397

Query Match 1.4% Score 8; DB 2; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 5.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 WMDPETET 414  
 |||||  
 Db 161 WMDPETET 168

## RESULT 8

S27707  
 daunorubicin resistance protein - Streptomyces peucetius  
 C;Species: Streptomyces peucetius  
 C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 02-Feb-2001  
 C;Accession: S27707  
 R;Guillouffe, P.G.; Hutchinson, C.R.  
 submitted to the EMBL Data Library, August 1991  
 A;Description: A bacterial analog of the mdr gene of mammalian tumor cells is present  
 A;Reference number: S27706  
 A;Accession: S27707  
 A;Molecule type: DNA  
 A;Residues: 1-330 <GUI>  
 A;Cross-references: EMBL:M73758; NID:g153228; PIDN:AAA74717.1; PID:g153230  
 C;Genetics:  
 A;Gene: drrA  
 A;Start codon: GTG

C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog  
 C;Keywords: antibiotic resistance; ATP; nucleotide binding; P-loop  
 F;24-215/Domain: ATP-binding cassette homology <ABC>  
 F;41-48/Region: nucleotide-binding motif A (P-loop)

Query Match 1.4% Score 8; DB 2; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 6.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 EVRSFSLG 66  
 |||||  
 Db 292 EVRSFSLG 299

## RESULT 9

T27762  
 hypothetical protein ZK177.6 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T27762  
 R;Anderson, K.

submitted to the EMBL Data Library, July 1995  
 A;Description: The sequence of C. elegans cosmid ZK177.  
 A;Reference number: Z20416  
 A;Accession: T27762  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-336 <AND>  
 A;Cross-references: EMBL:U21321; PIDN:AAB36970.1; GSPDB:GN00020; CESP:ZK177.6

A:Experimental source: strain Bristol N2; clone ZK177  
 C:Genetics:  
 A:Gene: CESP:ZK177.6  
 A:Map position: 2  
 A:Introns: 44/1; 75/3; 103/1; 185/3; 283/1

Query Match 1.4%; Score 8; DB 2; Length 336;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 525 DGTVKLWD 532  
 |||||  
 DB 219 DGTVKLWD 226

RESULT 10  
 I45774  
 odorant receptor 1 - channel catfish  
 C:Species: Ictalurus punctatus (channel catfish)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 04-Sep-1998  
 C:Accession: I45774  
 R:Ngai, J.; Dowling, M.M.; Buck, L.; Axel, R.; Chess, A.  
 Cell 72, 657-666, 1993  
 A:Title: The family of genes encoding odorant receptors in the channel catfish.  
 A:Reference number: A45774; MUID:93201590  
 A:Accession: I45774  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-344 <NGA>  
 A:Experimental source: olfactory epithelium  
 A:Note: sequence extracted from NCBI backbone (NCBIP:127744)  
 C:Superfamily: olfactory receptor OR14  
 C:Keywords: olfaction; transmembrane protein

Query Match 1.4%; Score 8; DB 2; Length 344;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IFSISIIA 40  
 |||||  
 DB 157 IFSISIIA 164

RESULT 11  
 T26805  
 hypothetical protein Y41C4A.11 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T26805  
 R:Steward, C.  
 submitted to the EMBL Data Library, October 1998  
 A:Reference number: Z20269  
 A:Accession: T26805  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-386 <WIL>  
 A:Cross-references: EMBL:AL032627; PIDN:CAA21549.1; CESP:Y41C4A.11  
 A:Experimental source: clone Y41C4A  
 C:Genetics:  
 A:Gene: CESP:Y41C4A.11  
 A:Introns: 24/3; 40/3; 100/1; 146/3; 306/3; 350/2

Query Match 1.4%; Score 8; DB 2; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 RIVSGSDD 282  
 |||||  
 DB 290 RIVSGSDD 297

RESULT 12  
 A43289  
 CDC4L protein - human (fragment)  
 N:Alternate names: cell division control protein CDC4 homolog  
 C:Species: Homo sapiens (man)  
 C:Date: 27-Apr-1993 #sequence\_revision 26-Jul-1996 #text\_change 15-Aug-1997  
 C:Accession: A43289; S30330  
 R:Feuchter, A.E.; Freeman, J.D.; Mager, D.L.  
 Genomics 13, 1237-1246, 1992  
 A:Title: Strategy for detecting cellular transcripts promoted by human endogenous lon  
 A:Reference number: A43289; MUID:92372019  
 A:Accession: A43289  
 A:Molecule type: mRNA  
 A:Residues: 1-515 <FEU>  
 A:Cross-references: EMBL:M83822  
 A:Note: sequence extracted from NCBI backbone (NCBIP:111997)  
 C:Genetics:  
 A:Gene: GDB:CDG4L  
 A:Cross-references: GDB:131404  
 A:Map position: 19p13.3-19p13.3

Query Match 1.4%; Score 8; DB 2; Length 515;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 SGRSDATL 365  
 |||||  
 DB 482 SGRSDATL 489

RESULT 13  
 T16607  
 hypothetical protein K10B2.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16607  
 R:Miller, N.  
 submitted to the EMBL Data Library, June 1995  
 A:Description: The sequence of C. elegans cosmid K10B2.  
 A:Reference number: Z18545  
 A:Accession: T16607  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-701 <MIL>  
 A:Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:K10B2.1  
 A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 1.4%; Score 8; DB 2; Length 701;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 446 RVWDVETG 453  
 |||||  
 DB 321 RVWDVETG 328

RESULT 14  
 A86243  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: A86243  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: A86243  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-821 <STO>  
A:Cross-references: GB:AE005172; NID:g1931650; PIDN:AAB65485.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 1.4%; Score 8; DB 2; Length 821;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 ITSSDDGT 527  
|||||||  
Db 426 ITSSDDGT 433

## RESULT 15

S56768  
capsid protein - Leishmania RNA virus 2 (fragment)  
C:Species: Leishmania RNA virus 2  
C:Date: 19-Mar-1997 #sequence\_revision 18-Jul-1997 #text\_change 18-Jul-1997  
C:Accession: S56768  
R:Widmer, G.; Dooley, S.  
Nucleic Acids Res. 23, 2300-2304, 1995  
A:Title: Phylogenetic analysis of Leishmania RNA virus and Leishmania suggests ancient v  
A:Reference number: S56768; MUID:95334386  
A:Accession: S56768  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-40 <WID>  
A:Cross-references: EMBL:L41164

Query Match 1.2%; Score 7; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 GRRVWSG 399  
|||||||  
Db 11 GRRVWSG 17

Search completed: September 25, 2001, 14:46:28  
Job time: 138 sec

**THIS PAGE BLANK (USPTO)**



```

Query Match          2.0%; Score 12; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 SGLSLTSIRVMD 449
DB 427 SGLSLTSIRVMD 438
      |||||
      |||||

RESULT 2
GROG_BOVIN
ID GROG_BOVIN STANDARD; PRT; 98 AA.
AC O46675;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GROWTH REGULATED PROTEIN HOMOLOG GAMMA PRECURSOR (GRO-GAMMA).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshimura T., Modi W.S.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U95811; AAB93927.1; -
CC HSPF; P09341; IMG5.
CC InterPro; IPR001089; -
CC InterPro; IPR001811; -
CC InterPro; IPR002473; -
CC Pfam; PF00048; IL8; 1.
CC PRINTS; PR00436; INTERLEUKIN8.
CC PRINTS; PR00437; SMALLCYTCKCX.
CC PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U95811; AAB93927.1; -
CC InterPro; IPR001089; -
CC InterPro; IPR001811; -
CC Pfam; PF00048; IL8; 1.
CC PRINTS; PR00437; SMALLCYTCKCX.
CC PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
CC Cytokine; Growth factor; Inflammatory response; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 98 GROWTH REGULATED PROTEIN HOMOLOG GAMMA.
FT DISULFID 39 65 BY SIMILARITY.
FT DISULFID 41 81 BY SIMILARITY.
SQ SEQUENCE 98 AA; 10393 MW; 942CD6897C21EDE9 CRC64;

Query Match          1.4%; Score 8; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQ 501
DB 40 QCLQTLQ 47
      |||||
      |||||

RESULT 3
GRO2_RABIT
ID GRO2_RABIT STANDARD; PRT; 104 AA.
AC P47854;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE GROWTH REGULATED PROTEIN HOMOLOG PRECURSOR (GRO HOMOLOG).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;

```

```

[1]
RN SEQUENCE FROM N.A.
RP Schwartz D., Chaverri-Alamada L., Berliner J., Kirchgessner T.,
RA Quisnonoro D., Fang J., Tekamp-Olson P., Luis J., Fogelman A.,
RA Territo M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A ROLE IN MONOCYTE ADHESION TO THE ENDOTHELIAL.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12310; AAA20487.1; -
CC HSPF; P09341; IMG5.
CC InterPro; IPR001089; -
CC InterPro; IPR001811; -
CC InterPro; IPR002473; -
CC Pfam; PF00048; IL8; 1.
CC PRINTS; PR00436; INTERLEUKIN8.
CC PRINTS; PR00437; SMALLCYTCKCX.
CC PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
CC Cytokine; Growth factor; Inflammatory response; Signal.
FT SIGNAL 1 31 PROBABLE.
FT CHAIN 32 104 GROWTH REGULATED PROTEIN HOMOLOG.
FT DISULFID 40 66 BY SIMILARITY.
FT DISULFID 42 82 BY SIMILARITY.
SQ SEQUENCE 104 AA; 10900 MW; 10B9D07B65C77F67 CRC64;

Query Match          1.4%; Score 8; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLQ 501
DB 41 QCLQTLQ 48
      |||||
      |||||

RESULT 4
GROA_BOVIN
ID GROA_BOVIN STANDARD; PRT; 104 AA.
AC O46676;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GROWTH REGULATED PROTEIN HOMOLOG ALPHA PRECURSOR (GRO-ALPHA).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshimura T., Modi W.S.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U95812; AAB93928.1; -

```



DR InterPro: IPR001089; -  
 DR InterPro: IPR001811; -  
 DR InterPro: IPR002473; -  
 DR Pfam: PF00048; IL8; 1.  
 DR PRINTS: PR00436; INTERLEUKIN8.  
 DR PRINTS: PR00437; SMALLCYTOKXC.  
 DR PROSITE: PS00471; SMALL\_CYTOKINES\_CXC; 1.  
 KW Cytokine; Growth factor; Inflammatory response; Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 104 GROWTH REGULATED PROTEIN HOMOLOG ALPHA.  
 FT DISULFID 40 66 BY SIMILARITY.  
 FT DISULFID 42 82 BY SIMILARITY.  
 SQ SEQUENCE 104 AA; 10984 MW; 1002CAC064DB1F76 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501  
 |||||  
 Db 41 QCLQTLOG 48

RESULT 5  
 GROB\_BOVIN STANDARD; PRT; 104 AA.  
 ID GROB\_BOVIN STANDARD; PRT; 104 AA.  
 AC O46677;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GROWTH REGULATED PROTEIN HOMOLOG BETA PRECURSOR (GRO-BETA).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yoshimura T., Modi W.S.;  
 Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE  
 C-X-C) (CHEMOKINE CXC).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U95813; AAB93929.1; -  
 DR InterPro: IPR001089; -  
 DR InterPro: IPR001811; -  
 DR InterPro: IPR002473; -  
 DR Pfam: PF00048; IL8; 1.  
 DR PRINTS: PR00436; INTERLEUKIN8.  
 DR PRINTS: PR00437; SMALLCYTOKXC.  
 DR PROSITE: PS00471; SMALL\_CYTOKINES\_CXC; 1.  
 KW Cytokine; Growth factor; Inflammatory response; Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 104 GROWTH REGULATED PROTEIN HOMOLOG BETA.  
 FT DISULFID 40 66 BY SIMILARITY.  
 FT DISULFID 42 82 BY SIMILARITY.  
 SQ SEQUENCE 104 AA; 10950 MW; 40A8C06A64D67F7B CRC64;

Query Match 1.4%; Score 8; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501

Db 41 QCLQTLOG 48  
 |||||  
 RESULT 6  
 GRO\_HUMAN STANDARD; PRT; 107 AA.  
 ID GRO\_HUMAN STANDARD; PRT; 107 AA.  
 AC P09341;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GROWTH REGULATED PROTEIN PRECURSOR (MELANOMA GROWTH STIMULATORY  
 DE ACTIVITY) (MGSA) (NEUTROPHIL-ACTIVATING PROTEIN 3) (NAP-3).  
 GN SCYB1 OR GRO1 OR GROA OR GRO OR MGSA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88041072; PubMed=2890161;  
 RA Anisowicz A., Bardwell L., Sager R.;  
 "Constitutive overexpression of a growth-regulated gene in  
 RT transformed Chinese hamster and human cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7188-7192(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88328991; PubMed=2970963;  
 RA Richmond A., Balentien E., Thomas H.G., Flagg G., Barton D.E.,  
 RA Spiess J., Bordon R., Francke U., Derynck R.;  
 "Molecular characterization and chromosomal mapping of melanoma  
 RT growth stimulatory activity, a growth factor structurally related to  
 RT beta-thromboglobulin.";  
 RL EMBO J. 7:2025-2033(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=91057157; PubMed=21295556;  
 RA Baker N.E., Kucera G., Richmond A.;  
 "Nucleotide sequence of the human melanoma growth stimulatory  
 RT activity (MGSA) gene.";  
 RL Nucleic Acids Res. 18:6453-6453(1990).  
 RN [4]  
 RP SEQUENCE OF 35-65.  
 RX MEDLINE=90217938; PubMed=2182761;  
 RA Schroeder J.-M., Persoon N.L.M., Christophers E.;  
 "Lipopolysaccharide-stimulated human monocytes secrete, apart from  
 RT neutrophil-activating peptide 1/interleukin 8, a second neutrophil-  
 RT activating protein. NH2-terminal amino acid sequence identity with  
 RT melanoma growth stimulatory activity.";  
 RL J. Exp. Med. 171:1091-1100(1990).  
 RN [5]  
 RP SEQUENCE OF 35-57.  
 RX MEDLINE=89246368; PubMed=2655583;  
 RA Golds E.E., Mason P., Nyirkos P.;  
 "Inflammatory cytokines induce synthesis and secretion of gro protein  
 RT and a neutrophil chemotactic factor but not beta 2-microglobulin in  
 RT human synovial cells and fibroblasts.";  
 RL Biochem. J. 259:585-588(1989).  
 RN [6]  
 RP POSSIBLE FUNCTION.  
 RX MEDLINE=89356650; PubMed=2670560;  
 RA Wen D., Rowland A., Derynck R.;  
 "Expression and secretion of gro/MGSA by stimulated human endothelial  
 RT cells.";  
 RL EMBO J. 8:1761-1766(1989).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=93387459; PubMed=8397104;  
 RA Fairbrother W.J., Reilly D., Colby T., Horuk R.;  
 "1H assignment and secondary structure determination of human  
 RT melanoma growth stimulating activity (MGSA) by NMR spectroscopy.";  
 RL FEBS Lett. 330:302-306(1993).

```

[8]
RP STRUCTURE BY NMR.
RX MEDLINE=94376296; PubMed=8089846;
RA Fairbrother W.J., Reilly D., Colby T., Hesselgesser J., Horuk R.;
RL "The solution structure of melanoma growth stimulating activity.";
RL J. Mol. Biol. 242:252-270(1994).
[9]
RP STRUCTURE BY NMR.
RX MEDLINE=95105175; PubMed=7806518;
RA Kim K.S., Clark-Lewis I., Sykes B.W.;
RL "Solution structure of GRO/melanoma growth stimulatory activity
determined by 1H NMR spectroscopy.";
RL J. Biol. Chem. 269:32909-32915(1994).
CC -1- FUNCTION: HAS CHEMOTACTIC ACTIVITY FOR NEUTROPHILS. MAY PLAY A
CC ROLE IN INFLAMMATORY AND EXERTS ITS EFFECTS ON ENDOTHELIAL CELLS
CC IN AN AUTOCRINE FASHION.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CXCL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03561; AAA35933.1; -
DR EMBL: X12510; CAA31027.1; -
DR EMBL: X34489; CAA38361.1; -
DR PIR: A28414; A28414.
DR PIR: S00983; S00983.
DR PIR: S03976; S03976.
DR PIR: S13669; S13669.
DR PDB: 1MGS; 30-SEP-94.
DR PDB: 1MSG; 31-MAR-95.
DR PDB: 1MSH; 31-MAR-95.
DR MIM: 155730; -
DR InterPro: IPR001089; -
DR InterPro: IPR001811; -
DR InterPro: IPR002473; -
DR Pfam: PF00048; IL8; 1.
DR PRINTS: PR00436; INTERLEUKIN8.
DR PRINTS: PR00437; SMALLCYTCKXC.
DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Growth factor; Inflammatory response; Signal; 3D-structure.
FT SIGNAL 1 34
FT CHAIN 35 107 GROWTH REGULATED PROTEIN.
FT DISULFID 43 69
FT DISULFID 45 85
SQ SEQUENCE 107 AA; 11301 MW; 17048A6B4D765CA2 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501
DB 44 QCLQTLOG 51

RESULT 7
M12A_HUMAN
ID M12A_HUMAN STANDARD; PRT; 107 AA.
AC P19875; Q9UPB8;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MACROPHAGE INFLAMMATORY PROTEIN-2-ALPHA PRECURSOR (MIP2-ALPHA) (GROWTH
DE REGULATED PROTEIN BETA) (GRO-BETA).
GN GRO2 OR GROB OR MIP2A.
OS Homo sapiens (Human).

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI\_TaxID=9606; [1] SEQUENCE FROM N.A. TISSUE=Histiocytic lymphoma; MEDLINE=90354792; PubMed=2201751; Tekamp-Olson P., Gallegos C., Bauer D., McClain J., Sherry B., Fabre M., van Deventer S., Cerami A.; "Cloning and characterization of cDNAs for murine macrophage inflammatory protein 2 and its human homologues"; J. Exp. Med. 172:911-919(1990). [2] SEQUENCE FROM N.A. MEDLINE=90377259; PubMed=2078213; Hida N., Grotendorst G.R.; "Cloning and sequencing of a new gro transcript from activated human monocytes: expression in leukocytes and wound tissue."; Mol. Cell. Biol. 10:5596-5599(1990). [3] SEQUENCE FROM N.A. MEDLINE=91017578; PubMed=2217207; Haskill S., Peace A., Morris J., Sporn S.A., Anisowicz A., Lee S.W., Smith T., Martin G., Ralph P., Sager R.; "Identification of three related human GRO genes encoding cytokine functions."; Proc. Natl. Acad. Sci. U.S.A. 87:7732-7736(1990). [4] SEQUENCE OF 35-107 FROM N.A. Jang J.S., Kim B.E.; Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases. [5] STRUCTURE BY NMR OF 39-107. MEDLINE=20069929; PubMed=10600366; Qian Y.Q., Johanson K.O., McDevitt P.; "Nuclear magnetic resonance solution structure of truncated human GRObeta [5-73] and its structural comparison with CXCR chemokine family members GROalpha and IL-8"; J. Mol. Biol. 294:1065-1072(1999). CC -1- FUNCTION: PRODUCED BY ACTIVATED MONOCYTES AND NEUTROPHILS AND EXPRESSED AT SITES OF INFLAMMATION. CC -1- SIMILARITY: BELONGS TO THE INTERCERIN ALPHA FAMILY (SMALL CYTOKINE C-X-C) (CHEMOKINE CXCL). CC ----- CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)). CC ----- CC EMBL: X53799; CAA37808.1; - EMBL: M36820; AAA63183.1; - EMBL: M57731; AAA63182.1; - EMBL: AF043340; AAC03540.1; - PIR: JH0281; JH0281. PDB: 1QNK; 04-FEB-00. MIM: 139110; - InterPro: IPR001089; - InterPro: IPR001811; - InterPro: IPR002473; - Pfam: PF00048; IL8; 1. PRINTS: PR00436; INTERLEUKIN8. PRINTS: PR00437; SMALLCYTOKINE. PROSITE: PS00471; SMALL\_CYTOKINES\_CXC; 1. Cytokine; Chemotaxis; Inflammatory response; Signal; 3D-structure. SIGNAL 1 34 FT CHAIN 35 107 MACROPHAGE INFLAMMATORY PROTEIN-2-ALPHA. FT FT DISULFID 43 69 FT FT DISULFID 45 85 FT SEQUENCE 107 AA; 11389 MW; 740F277E928571BA CRC64; SO

```

Query Match          1.4%  Score 8;  DB 1;  Length 107;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 QCLQTLOG 501
DB 44 QCLQTLOG 51

RESULT 8
MI2B_HUMAN          STANDARD;      PRT;  107 AA.
AC P19876;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MACROPHAGE INFLAMMATORY PROTEIN-2-BETA PRECURSOR (MIP2-BETA) (GROWTH
DE REGULATED PROTEIN GAMMA) (GRO-GAMMA).
GN GRO3 OR GROG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Histiocytic lymphoma;
RX MEDLINE=90354792; PubMed=2201751;
RA Tekamp-Olson P., Gallagos C., Bauer D., McClain J., Sherry B.,
RA Fabre M., van Deventer S., Cerami A.;
RT "Cloning and characterization of cDNAs for murine macrophage
RT inflammatory protein 2 and its human homologues.";
RL J. Exp. Med. 172:911-919(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91017578; PubMed=2217207;
RA Haskill S., Peace A., Morris J., Sporn S.A., Anisowicz A., Lee S.W.,
RA Smith T., Martin G., Ralph P., Sager R.;
RT "Identification of three related human GRO genes encoding cytokine
RT functions.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:7732-7736(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN INFLAMMATION AND EXERT ITS EFFECTS
CC ON ENDOTHELIAL CELLS IN AN AUTOCRINE FASHION.
CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
CC C-X-C) (CHEMOKINE CX-C).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53800; CAA37809.1; -
DR EMBL; M36821; AAA63184.1; -
DR PIR; B38290; B38290.
DR PIR; JH0282; JH0282.
DR HSP; P09341; 1MGS.
DR MIM; 139111; -
DR InterPro; IPR001089; -
DR InterPro; IPR001811; -
DR InterPro; IPR002473; -
DR Pfam; PF00048; IL8; 1.
DR PRINTS; PR00436; INTERLEUKIN8.
DR PRINTS; PR00437; SMALLCYTCKXC.
DR PROSITE; PS00471; SMALL_CYTOKINES_CXC; 1.
KW Cytokine; Chemotaxis; Inflammatory response; Signal.
FT SIGNAL 1 34
FT CHAIN 35 107 MACROPHAGE INFLAMMATORY PROTEIN-
FT 2-BETA.
FT DISULFID 43 69 BY SIMILARITY.

Query Match          1.4%  Score 8;  DB 1;  Length 330;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 EVRSFSLG 66
DB 292 EVRSFSLG 299

RESULT 10
YSSL_CAEEL          STANDARD;      PRT;  701 AA.
ID YSSL_CAEEL
AC Q09990;
DT 15-JUL-1998 (Rel. 36, Created)

```

15-JUL-1998 (Rel. 36, Last sequence update)  
 15-JUL-1998 (Rel. 36, Last annotation update)  
 HYPOTHEICAL 80.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN K10B2.1 IN  
 CHROMOSOME II.  
 GN Caenorhabditis elegans.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Miller N.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: STRONG, TO X.LAEVIS BETA-TRCP.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U28730; AAA68258.1; -.  
 DR WormPep; K10B2.1; C802008.  
 DR InterPro; IPR001680; -.  
 DR Pfam; PF00400; WD40; 7.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR PROSITE; PS0181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS.1; 5.  
 DR PROSITE; PS0082; WD\_REPEATS.2; 7.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 FT REPEAT 256 284 WD 1.  
 FT REPEAT 296 324 WD 2.  
 FT REPEAT 336 364 WD 3.  
 FT REPEAT 379 407 WD 4.  
 FT REPEAT 419 447 WD 5.  
 FT REPEAT 459 487 WD 6.  
 FT REPEAT 508 536 WD 7.  
 FT DOMAIN 606 615 POLY-ALA.  
 SQ SEQUENCE 701 AA; 80320 MW; 69FA0B00F83270E3 CRC64;  
 Query Match 1.4%; Score 8; DB 1; Length 701;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 446 RWMVDVETG 453  
 Db 321 RWMVDVETG 328  
 RESULT 11  
 CC4H\_HUMAN  
 ID CC4H\_HUMAN STANDARD; PRT; 1918 AA.  
 AC P50851;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CDC4-LIKE PROTEIN (BEIGE-LIKE PROTEIN) (FRAGMENT).  
 GN CDC4L OR BGL.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mager D.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RH [2]

RP SEQUENCE OF 1405-1918 FROM N.A.  
 RX MEDLINE=92372019; PubMed=1505956;  
 RA Feuchter A.E.; Freeman J.D.; Mager D.L.;  
 RT "Strategy for detecting cellular transcripts promoted by human  
 RT endogenous long terminal repeats: identification of a novel gene  
 RT (CDC4L) with homology to yeast CDC4.";  
 RL Genomics 13:1237-1246(1992).  
 CC -!- SIMILARITY: CONTAINS AT LEAST 2 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: CONTAINS 1 BEACH DOMAIN.  
 CC -!- CAUTION: WAS SAID TO BE SIMILAR TO YEAST CDC4, BUT THAT SIMILARITY  
 CC IS VERY LIMITED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M83822; AAB09603.1; -.  
 DR InterPro; IPR001680; -.  
 DR Pfam; PF00400; WD40; 2.  
 DR PROSITE; PS0197; BEACH; 1.  
 DR PROSITE; PS00678; WD\_REPEATS.1; FALSE\_NEG.  
 DR PROSITE; PS0082; WD\_REPEATS.2; FALSE\_NEG.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; FALSE\_NEG.  
 KW Repeat; WD repeat.  
 FT NON\_TER 1 1  
 FT DOMAIN 1445 1715 BEACH.  
 FT REPEAT 1817 1859 WD 2.  
 FT REPEAT 1862 1905 WD 3.  
 SQ SEQUENCE 1918 AA; 213598 MW; 427EA619BDF9A724 CRC64;  
 Query Match 1.4%; Score 8; DB 1; Length 1918;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 358 SGRSDATL 365  
 Db 1885 SGRSDATL 1892  
 RESULT 12  
 FERS\_MAIZE  
 ID FERS\_MAIZE STANDARD; PRT; 135 AA.  
 AC P27789;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE FERREDOXIN V PRECURSOR (FD V).  
 GN FDV5 OR PFD5.  
 OS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
 CC Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hase T.; Kimatsa Y.; Yonekura K.; Matsumura T.; Sakakibara H.;  
 RT "Molecular cloning and differential expression of the maize ferredoxin  
 RT gene family.";  
 RL Plant Physiol. 96:77-83(1991).  
 CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -!- COPACTOR: BINDS A 2FE-2S CLUSTER.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M73828; AAA33462.1; -.  
 DR HSSP: P27320; IDOY.  
 DR MaizeDB: 66392; -.  
 DR InterPro: IPR000564; -.  
 DR InterPro: IPR001041; -.  
 DR Pfam: PF00111; fer2; 1.  
 DR PRINTS: PR00159; 2FE2SFRDXIN.  
 DR PROSITE: PS00197; 2FE2S-FERREDOXIN; 1.  
 KW Electron transport; Iron-sulfur; Chloroplast; Transit peptide;  
 KW Multigene family.  
 FT TRANSIT 1 38 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 39 135 FERREDOXIN V.  
 FT METAL 77 77 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 82 82 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 85 85 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 SQ SEQUENCE 135 AA; 14399 MW; 8FA43C41AD4CB976 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 VVSGSLD 442

||||||

DB 89 VVSGSLD 95

#### RESULT 13

ID DCBI\_METFR STANDARD; PRT; 170 AA.  
 AC Q49162;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CARBON MONOXIDE DEHYDROGENASE I BETA SUBUNIT (EC 1.2.99.2).  
 GN CDH1B.  
 OS Methanosarcina frisia.  
 OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;  
 OC Methanosarcina.  
 OX NCBI\_TaxID=2212;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GOEL.  
 RX MEDLINE=96278885; PubMed=8662887;  
 RA Eggen R.I.L., van Kranenburg R., Vriesema A.J.M., Geerling A.C.M.,  
 RA Verhagen M.F.J.M., Hagen W.R., de Vos W.M.;  
 RT "Carbon monoxide dehydrogenase from Methanosarcina frisia Gol.  
 RT Characterization of the enzyme and the regulated expression of two  
 RT operon-like cdh gene clusters";  
 RL J. Biol. Chem. 271:14256-14263(1996).  
 CC -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED  
 CC ACCEPTOR.  
 CC -!- COFACTOR: NICKEL ION.  
 CC -!- SUBUNIT: HOMOTETRAMER OF TWO ALPHA AND TWO BETA CHAINS.  
 CC -!- MISCELLANEOUS: OPTIMUM ACTIVITY IS MEASURED BETWEEN PH 8 AND 9.  
 CC -!- SIMILARITY: TO M.SOHNGENII CARBON MONOXIDE DEHYDROGENASE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L26487; AAC37045.1; -.  
 KW Oxidoreductase; Nickel.  
 SQ SEQUENCE 170 AA; 18696 MW; 74B309E6487E3E79 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 ALDELID 140

||||||

DB 162 ALDELID 168

#### RESULT 14

UNG\_HSV2  
 ID UNG\_HSV2 STANDARD; PRT; 255 AA.  
 AC P53765;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE URACIL-DNA GLYCOSYLASE (EC 3.2.2.-) (UDG).  
 GN 46.  
 OS Equine herpesvirus type 2 (strain 86/87) (EHV-2).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 OX NCBI\_TaxID=82831;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95302501; PubMed=7783207;  
 RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus 2";  
 RL J. Mol. Biol. 249:520-528(1995).  
 CC -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE  
 CC AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA  
 CC POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE.  
 CC -!- SIMILARITY: BELONGS TO THE URACIL-DNA GLYCOSYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U20824; AAC13834.1; -.  
 DR HSSP: P13051; 1AKZ.  
 DR InterPro: IPR002043; -.  
 DR Pfam: PF00315; UNG; 1.  
 DR PROSITE: PS00130; U\_DNA\_GLYCOSYLASE; 1.  
 KW DNA repair; Hydrolase; Glycosidase.  
 FT ACT\_SITE 90 90 GENERAL BASE (BY SIMILARITY).  
 SQ SEQUENCE 255 AA; 29099 MW; 20104402C5297336 CRC64;

Query Match 1.2%; Score 7; DB 1; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 CEPTQVK 148

||||||

DB 77 CEPTQVK 83

#### RESULT 15

GBLP\_SCHPO STANDARD; PRT; 314 AA.  
 ID GBLP\_SCHPO  
 AC Q10281; P78896;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN.  
 GN SPAC6B12.15.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

```
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCHI_taxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ED616;
RA Park S.K., Yoo H.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 5-314 FROM N.A.
RC STRAIN-PR745;
RA Yoshioka S., Kato K., Okayama H.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L37885; AAA56865.1; -.
DR EMBL; Z98531; CAB11079.1; -.
DR EMBL; D89247; BAA13908.1; -.
DR InterPro; IPR001680; -.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS00082; WD_REPEATS_2; 6.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
RW Repeat; WD repeat.
FT REPEAT 13 44 WD 1.
FT REPEAT 61 91 WD 2.
FT REPEAT 103 133 WD 3.
FT REPEAT 146 178 WD 4.
FT REPEAT 190 220 WD 5.
FT REPEAT 231 260 WD 6.
FT REPEAT 281 311 WD 7.
FT CONFLICT 5 5 L -> I (IN REF. 2).
FT CONFLICT 7 7 L -> V (IN REF. 2).
FT CONFLICT 89 89 L -> W (IN REF. 2).
FT CONFLICT 137 137 N -> Y (IN REF. 2).
FT CONFLICT 168 168 A -> V (IN REF. 2).
FT CONFLICT 171 171 D -> E (IN REF. 2).
FT CONFLICT 173 173 A -> P (IN REF. 2).
FT CONFLICT 176 176 V -> A (IN REF. 2).
FT CONFLICT 244 244 N -> D (IN REF. 2).
SQ SEQUENCE 314 AA; 34851 MW; 4E14707164E68ACD CRC64;
```

```
Query Match 1.28; Score 7; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 356 VVSGSRD 362
Db 120 VVSGSRD 126
|||||||
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:45:36 ; Search time 22.42 seconds  
(without alignments)  
3475.810 Million cell updates/sec

Title: US-09-328-877A-8  
Perfect score: 589  
Sequence: 1 MSKPKPTLNHGLVPVDLKS.....SRNGTEETKLVLDFDVMK 589

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	540	91.7	553	4 Q9NUX6	Q9nuX6 homo sapien
2	21	3.6	1326	5 Q9VZF4	Q9vzf4 drosophila
3	12	2.0	587	5 O44083	O44083 caenorhabdi
4	10	1.7	310	4 O75938	O75938 homo sapien
5	10	1.7	357	4 O95320	O95320 homo sapien
6	9	1.5	472	4 Q9Y6J1	Q9y6j1 homo sapien
7	9	1.5	524	4 Q9NSU1	Q9nsul homo sapien
8	9	1.5	703	11 Q9WVM9	Q9wvm9 mus musculu
9	9	1.5	741	4 Q9H9K3	Q9h9k3 homo sapien
10	9	1.5	845	10 Q9LJ3	Q9ljr3 arabidopsia
11	9	1.5	1028	4 Q9HCM8	Q9hcm8 homo sapien
12	9	1.5	2904	11 Q9EPN0	Q9epn0 mus musculu
13	9	1.5	2931	11 Q9EPW9	Q9epw9 mus musculu
14	9	1.5	2936	11 Q9EPN1	Q9epn1 mus musculu
15	8	1.4	272	2 O66716	O66716 aquifex aeo
16	8	1.4	276	5 Q9VXP8	Q9vxp8 drosophila
17	8	1.4	326	5 Q9YXF7	Q9yxf7 drosophila
18	8	1.4	326	5 Q9VAJ2	Q9vaj2 drosophila
19	8	1.4	327	5 O76523	O76523 drosophila

20	8	1.4	337	4	Q9H2U3	Q9h2u3 homo sapien
21	8	1.4	344	13	Q9PSJ1	Q9psj1 ictalurus p
22	8	1.4	386	5	Q9XWU3	Q9xwu3 caenorhabdi
23	8	1.4	476	5	Q9VG17	Q9vg17 drosophila
24	8	1.4	507	5	Q09661	Q09661 caenorhabdi
25	8	1.4	543	5	Q9V6J3	Q9v6j3 drosophila
26	8	1.4	665	5	Q9GNN6	Q9gnn6 caenorhabdi
27	8	1.4	821	10	O04093	O04093 arabidopsia
28	8	1.4	2792	11	Q9ESD4	Q9esd4 mus musculu
29	8	1.4	2956	11	Q9ESE1	Q9esel mus musculu
30	7	1.2	41	14	Q83070	Q83070 leishmania
31	7	1.2	41	14	Q83098	Q83098 leishmania
32	7	1.2	44	14	Q9QEI2	Q9qe12 human immun
33	7	1.2	52	2	O33534	O33534 rhizobium l
34	7	1.2	123	11	O61621	O61621 mus musculu
35	7	1.2	139	2	O46367	O46367 chlorobium
36	7	1.2	139	2	O46466	O46466 chlorobium
37	7	1.2	149	5	Q21091	Q21091 caenorhabdi
38	7	1.2	150	2	Q9WMX4	Q9wxw4 pseudomonas
39	7	1.2	150	5	O06784	O06784 haematobia
40	7	1.2	163	2	O9X5U1	O9x5ul streptomyce
41	7	1.2	164	5	O9TYL1	O9tyl1 caenorhabdi
42	7	1.2	170	1	P72020	P72020 methanosarc
43	7	1.2	176	5	Q9GV52	Q9gv52 drosophila
44	7	1.2	187	1	Q9YF07	Q9yf07 aeropyrum p
45	7	1.2	190	2	O86782	O86782 streptomyce

#### ALIGNMENTS

RESULT 1  
Q9NUX6 PRELIMINARY; PRT; 553 AA.  
ID Q9NUX6  
AC Q9NUX6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE CDNA FLJ11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10  
DE PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. -  
RC TISSUE=PLACENTA;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki Y., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK001933; BAA91986.1;  
DR InterPro: IPR001680;  
DR InterPro: IPR001810;  
DR Pfam: PF00400; WD40; 7;  
DR Pfam: PF00646; F-box; 1.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS00181; FBOX; 1.  
DR PROSITE; PS00678; WD.REPEATS; UNKNOWN\_5.  
DR SMART; SM00256; FBOX; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 553 AA; 62280 MW; CA829C221986A3F2 CRC64;

Query Match 91.7%; Score 540; DB 4; Length 553;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 MKRKLHGSEVRSFSGKPKCKVSEYTSITGLVPCSATPTTGLDLRAANGQQRRTTS 109

DB 14 MKRKLHGSEVRSFSGKPKCKVSEYTSITGLVPCSATPTTGLDLRAANGQQRRTTS 73

QY 110 VQPPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQDFISLLPKE 169  
DQ 74 VQPPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQDFISLLPKE 133  
QY 170 LALYVLSLEPKDLQAQTCRYWRIAEADNLLWREKCEEGIDEPHLKRRKVIKPGFI 229  
DQ 134 LALYVLSLEPKDLQAQTCRYWRIAEADNLLWREKCEEGIDEPHLKRRKVIKPGFI 193  
QY 230 HSPWKSAYIRQHRIDTNRGELKSPKVLKHDDHVVITCLQFCGRIYVSGSDNTLTKVWS 289  
DQ 194 HSPWKSAYIRQHRIDTNRGELKSPKVLKHDDHVVITCLQFCGRIYVSGSDNTLTKVWS 253  
QY 290 AVTGKCLRTLVTGHTGVWSSQMRDNIISGSDTRTLKVNNAETGECIHTLTCGHTSTVRCM 349  
DQ 254 AVTGKCLRTLVTGHTGVWSSQMRDNIISGSDTRTLKVNNAETGECIHTLTCGHTSTVRCM 313  
QY 350 HLHKKRVVSGSRDATLVRWDIETGQCLHVLGMHVAARCVQYDGRVRYVSGAYDFMVKVWD 409  
DQ 314 HLHKKRVVSGSRDATLVRWDIETGQCLHVLGMHVAARCVQYDGRVRYVSGAYDFMVKVWD 373  
QY 410 PTEETCLHTLQHTNRVYSLQFDGHHVYVSGSLDTSIRVMDVETGNCIHTLTCGHTSTVRCM 469  
DQ 374 PTEETCLHTLQHTNRVYSLQFDGHHVYVSGSLDTSIRVMDVETGNCIHTLTCGHTSTVRCM 433  
QY 470 ELKONILVSGNADSVTKIWDIKTGQCLQTLQLOGPNKHQSAVTCLOFNKKNFVITSSDDGTVK 529  
DQ 434 ELKONILVSGNADSVTKIWDIKTGQCLQTLQLOGPNKHQSAVTCLOFNKKNFVITSSDDGTVK 493  
QY 530 LWDLKTGFIRNLVLESGGGVGVWRIRASNTKLVCVAVSGRNGTEETKLLVLDVDFDVKM 589  
DQ 494 LWDLKTGFIRNLVLESGGGVGVWRIRASNTKLVCVAVSGRNGTEETKLLVLDVDFDVKM 553  
RESULT 2  
QYVZF4 QYVZF4  
ID Q9VZF4 PRELIMINARY; PRT; 1326 AA.  
AC Q9VZF4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE CG15010 PROTEIN.  
GN CG15010.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA Sutton G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazew R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahle C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.B.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Relneit K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003480; AAG2246.1; -.  
DR FlyBase; FBgn0035516; CG15010.  
DR InterPro; IPR001680; -.  
DR InterPro; IPR001810; -.  
DR Pfam; PF00400; WD40; 7.  
DR Pfam; PF00646; F-box; 1.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS0181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS; 5.  
KW Repeat; WD repeat.  
SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;  
Query Match 3.6%; Score 21; DB 5; Length 1326;  
Best Local Similarity 100.0%; Pred. No. 1.9e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 435 HVVSGSLDTSIRVMDVETGNC 455  
DQ 1167 HVVSGSLDTSIRVMDVETGNC 1187  
RESULT 3  
QYVZF4 QYVZF4  
ID Q44083 PRELIMINARY; PRT; 587 AA.  
AC Q44083;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE SEL-10.  
GN SEL-10.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98051191; PubMed=9389650;  
RA Hubbard E.J.A., Wu G., Kitajewski J., Greenwald I.;  
RT "sel-10, a negative regulator of lin-12 activity in Caenorhabditis  
elegans, encodes a member of the CDC4 family of proteins.";  
RL Genes Dev. 11:3182-3193(1997).  
DR EMBL; AF020788; AAC47809.1; -.  
DR InterPro; IPR000412; -.  
DR InterPro; IPR001680; -.  
DR InterPro; IPR001810; -.  
DR Pfam; PF00400; WD40; 7.  
DR Pfam; PF00646; F-box; 1.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS00890; ABC2\_MEMBRANE; UNKNOWN\_1.  
DR PROSITE; PS0181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS; UNKNOWN\_5.  
KW Repeat; WD repeat.  
SQ SEQUENCE 587 AA; 65311 MW; 2D3970B4EFAA1B8C CRC64;



Query Match 2.0%; Score 12; DB 5; Length 587;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 SGLDTSIRWMD 449  
|||||  
DB 435 SGLDTSIRWMD 446

RESULT 4  
075938 PRELIMINARY; PRT; 310 AA.  
AC 075938;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE 38 KDA SPLICING FACTOR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98400255; PubMed=9731529;  
RA Neubauer G., King A., Rappsilber J., Calvio C., Watson M., Ajuh P.,  
RT Sleeman J., Lamond A., Mann M.;  
RT "Mass spectrometry and EST-database searching allows characterization  
of the multi-protein spliceosome complex.";  
RL Nat. Genet. 20:46-50(1998).  
DR EMBL; AF083383; AAC64084.1; -.  
DR InterPro; IPR001680; -.  
DR InterPro; IPR002106; -.  
DR Pfam; PF00400; WD40; 6.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
DR PROSITE; PS00678; WD\_REPEATS; UNKNOWN\_5.  
DR SMART; SM00320; WD40; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 310 AA; 34290 MW; 7347F144E7D9796E CRC64;

Query Match 1.7%; Score 10; DB 4; Length 310;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 SDDGTVKLWD 532  
|||||  
DB 112 SDDGTVKLWD 121

RESULT 5  
095320 PRELIMINARY; PRT; 357 AA.  
AC 095320;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE U5 SNRNP-SPECIFIC 40 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98449962; PubMed=9774689;  
RA Achsel T., Ahrens K., Brahm H., Teigelkamp S., Luhrmann R.;  
RT "The human U5-220KD protein (hrp8) forms a stable RNA-free complex  
with several U5-specific proteins, including an RNA unwindase, a  
RT homologue of ribosomal elongation factor EF-2, and a novel WD-40  
protein.";  
RL Mol. Cell. Biol. 18:6756-6766(1998).  
DR EMBL; AF090988; AAC59625.1; -.  
DR InterPro; IPR001680; -.

DR InterPro; IPR002106; -.  
DR Pfam; PF00400; WD40; 7.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
DR PROSITE; PS00678; WD\_REPEATS; UNKNOWN\_5.  
DR SMART; SM00320; WD40; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 357 AA; 39298 MW; 87A265539437BD86 CRC64;

Query Match 1.7%; Score 10; DB 4; Length 357;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 SDDGTVKLWD 532  
|||||  
DB 171 SDDGTVKLWD 180

RESULT 6  
Q9Y6J1 PRELIMINARY; PRT; 472 AA.  
AC Q9Y6J1;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE LYOSOMAL TRAFFICKING REGULATOR 2 (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Tchernev V.T., McMurtrie E.B., Nguyen Q.A., Mishra V.S.,  
RA Barbosa M.D.F.S., McIndoe R., Kingmore S.F.;  
RT "Identification of LYST2, a brain-specific member of the Chediak-  
RT Higashi syndrome gene family.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF072371; AAD41633.1; -.  
DR InterPro; IPR000409; -.  
DR InterPro; IPR001680; -.  
DR Pfam; PF00400; WD40; 4.  
DR Pfam; PF02138; Beach; 1.  
DR PROSITE; PS50197; BEACH; 1.  
DR SMART; SM00320; WD40; 1.  
KW Repeat; WD repeat.  
FT NON\_TER 1  
SQ SEQUENCE 472 AA; 52810 MW; 679A710103FDAE5E CRC64;

Query Match 1.5%; Score 9; DB 4; Length 472;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 VSGSRDRTL 365  
|||||  
DB 313 VSGSRDRTL 321

RESULT 7  
Q9NSU1 PRELIMINARY; PRT; 524 AA.  
AC Q9NSU1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-WAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 58.8 KDA PROTEIN.  
GN DKF2P434G0511.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RA Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137748; CAB70903.1; -.
DR InterPro; IPR000409; -.
DR InterPro; IPR001680; -.
DR Pfam; PF00400; WD40; 5.
DR Pfam; PF02138; Beach; 2.
DR PROSITE; PS0197; BEACH; 1.
KW Hypothetical protein.
SQ SEQUENCE 524 AA; 58801 MW; F4A38C3EE58D390 CRC64;

Query Match 1.5%; Score 9; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 357 VSGSRDATL 365
Db 318 VSGSRDATL 326

RESULT 8
ID Q9WM9 PRELIMINARY; PRT; 703 AA.
AC Q9WM9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE LYSOSOMAL TRAFFICKING REGULATOR 2 (FRAGMENT).
GN LYST2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP Tchernev V.T., McMurtrie E.B., Nguyen Q.A., Mishra V.S.,
RA Barbosa M.D.F.S., McIndoe R., Kingsmore S.F.;
RT "Identification of LYST2, a brain-specific member of the Chediak-
RT Higashi syndrome gene family";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF072372; AAD41634.1; -.
DR InterPro; IPR000409; -.
DR InterPro; IPR001680; -.
DR Pfam; PF00400; WD40; 5.
DR Pfam; PF02138; Beach; 1.
DR PROSITE; PS0197; BEACH; 1.
DR SMART; SM00320; WD40; 1.
FT NON_TER 1
SQ SEQUENCE 703 AA; 79349 MW; 6776B57B1D972127 CRC64;

Query Match 1.5%; Score 9; DB 11; Length 703;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 357 VSGSRDATL 365
Db 511 VSGSRDATL 519

RESULT 9
ID Q9H9K3 PRELIMINARY; PRT; 741 AA.
AC Q9H9K3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CDNA FLJ12686 F15, CLONE NT2RM4002527, WEAKLY SIMILAR TO VEGETATIBLE
DE INCOMPATIBILITY PROTEIN HET-E-1 (FRAGMENT).
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022748; BAB14222.1; -.
FT NON_TER 741
SQ SEQUENCE 741 AA; 82214 MW; 8BE2038217996D07 CRC64;

Query Match 1.5%; Score 9; DB 4; Length 741;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 521 TSSDDGTVK 529
Db 84 TSSDDGTVK 92

RESULT 10
ID Q9LJR3 PRELIMINARY; PRT; 845 AA.
AC Q9LJR3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PHOTOMORPHOGENESIS REPRESSOR PROTEIN-LIKE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones";
RL DNA Res. 7:217-221(2000).
DR EMBL; AP000413; BAB02165.1; -.
DR InterPro; IPR00719; -.
DR InterPro; IPR001680; -.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_2.
DR SMART; SM00221; STYKC; 1.
KW ATP-binding; Repeat; Transferase; WD repeat.
SQ SEQUENCE 845 AA; 94353 MW; EB5ADEFD16F316 CRC64;

Query Match 1.5%; Score 9; DB 10; Length 845;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 523 SDDGTVKLW 531

```

Db 646 SDDGTVKWLW 654  
|||||

RESULT 11  
Q9HGM8 PRELIMINARY; PRT; 1028 AA.  
ID Q9HGM8  
AC Q9HGM8  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE KIAA1544 PROTEIN (FRAGMENT).  
GN KIAA1544.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=1097877;  
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes.  
RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
RT code for large proteins in vitro."  
RL DNA Res. 7:273-281(2000).  
DR EMBL; AB046764; BAB13370.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 1028 AA; 116708 MW; 1204AAA8AA03B60F CRC64;

Query Match 1.5%; Score 9; DB 4; Length 1028;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 357 VSGSRDRTL 365  
Db 822 VSGSRDRTL 830  
|||||

RESULT 12  
Q9EPN0 PRELIMINARY; PRT; 2904 AA.  
ID Q9EPN0  
AC Q9EPN0  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE NEUROBEACHIN.  
GN NBEA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=1102458;  
RA Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,  
RA Petrasch-Parwez E., Killmann M.W.;  
RT "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi  
RT protein homolog implicated in neuronal membrane traffic."  
RL J. Neurosci. 20:8551-8565(2000).  
DR EMBL; Y18276; CAC18812.1; -.  
SQ SEQUENCE 2904 AA; 323198 MW; B41D4ADF71AD6FF4 CRC64;

Query Match 1.5%; Score 9; DB 11; Length 2904;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 357 VSGSRDRTL 365  
Db 2698 VSGSRDRTL 2706  
|||||

RESULT 13  
Q9EPM9 PRELIMINARY; PRT; 2931 AA.  
ID Q9EPM9  
AC Q9EPM9  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE NEUROBEACHIN.  
GN NBEA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=11102458;  
RA Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,  
RA Petrasch-Parwez E., Killmann M.W.;  
RT "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi  
RT protein homolog implicated in neuronal membrane traffic."  
RL J. Neurosci. 20:8551-8565(2000).  
DR EMBL; Y18276; CAC18813.1; -.  
SQ SEQUENCE 2931 AA; 326149 MW; B2E2E8E87E6B2696 CRC64;

Query Match 1.5%; Score 9; DB 11; Length 2931;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 357 VSGSRDRTL 365  
Db 2725 VSGSRDRTL 2733  
|||||

RESULT 14  
Q9EPN1 PRELIMINARY; PRT; 2936 AA.  
ID Q9EPN1  
AC Q9EPN1  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE NEUROBEACHIN.  
GN NBEA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=11102458;  
RA Wang X., Herberg F.W., Laue M.M., Wullner C., Hu B.,  
RA Petrasch-Parwez E., Killmann M.W.;  
RT "Neurobeachin: A protein kinase A-anchoring, beige/Chediak-Higashi  
RT protein homolog implicated in neuronal membrane traffic."  
RL J. Neurosci. 20:8551-8565(2000).  
DR EMBL; Y18276; CAC18811.1; -.  
SQ SEQUENCE 2936 AA; 326738 MW; A6CFD90CA666CEA4 CRC64;

Query Match 1.5%; Score 9; DB 11; Length 2936;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 357 VSGSRDRTL 365  
Db 2730 VSGSRDRTL 2738  
|||||

RESULT 15  
O66716

ID O66716 PRELIMINARY; PRT; 272 AA.  
AC O66716;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE HYPOTHETICAL 31.6 KDA PROTEIN.  
GN AQ\_397.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID:63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE=98196666; PubMed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Shead M.A., Keller M., Aujay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus.";  
RL Nature 392:353-358(1998).  
DR EMBL; AE000688; AAC06680.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 272 AA; 31561 MW; 652D06CG2E15D581 CRC64;

Query Match 1.48; Score 8; DB 2; Length 272;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 407 VNDPETET 414  
| | | | | | | |  
Db 161 VNDPETET 168

Search completed: September 25, 2001, 14:49:07  
Job time: 211 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:38:25 ; Search time 20.43 seconds  
(without alignments)  
1747.799 Million cell updates/sec

Title: US-09-328-877A-8  
Perfect score: 3143  
Sequence: 1 MSKPGKPTLNHGLVPDLKS.....SRNGTEETKLLVDFDVKM 589

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3143	100.0	589	20 AAY22466	Human mammary sel-
2	3143	100.0	589	22 AAB59198	Human mammary sel-
3	3140	99.9	589	21 AAB01204	Human GTPase assoc
4	2983	94.9	559	20 AAY22467	Human mammary sel-
5	2983	94.9	559	22 AAB59199	Human mammary sel-
6	2890.5	92.0	627	20 AAY22461	Human hippocampal
7	2890.5	92.0	627	22 AAB59193	Human hippocampal
8	2890.5	92.0	666	20 AAY22471	Human C-term mychi
9	2890.5	92.0	666	22 AAB59203	C-terminal mychis
10	2890.5	92.0	669	20 AAY22470	Human Cterm V5 his
11	2890.5	92.0	669	22 AAB59202	C-terminal V5 His

12	2890	92.0	592	20 AAY22462	Human hippocampal
13	2890	92.0	592	22 AAB59194	Human hippocampal
14	2889	91.9	545	20 AAY22464	Human hippocampal
15	2889	91.9	545	22 AAB59196	Human hippocampal
16	2889	91.9	553	20 AAY22463	Human hippocampal
17	2889	91.9	553	22 AAB59195	Human hippocampal
18	2887	91.9	540	20 AAY22465	Human hippocampal
19	2887	91.9	540	22 AAY22468	Human mammary sel-
20	2887	91.9	540	22 AAB59197	Human hippocampal
21	2887	91.9	540	22 AAB59200	Human mammary sel-
22	2887	91.9	626	20 AAY22469	Human 6myc-N-sel-1
23	2887	91.9	626	22 AAB59201	Protein encoded by
24	1238	39.4	587	20 AAY03204	Amino acid sequenc
25	650	20.7	640	21 AAY83252	F-box protein Met3
26	641.5	20.4	542	21 AAY96896	Human E3 ubiquitin
27	641	20.4	703	21 AAY83253	A human beta-trans
28	640	20.4	569	20 AAY24054	Human beta-transdu
29	640	20.4	569	21 AAB12813	Human beta-T-CP
30	640	20.4	569	21 AAY96697	F-box protein FBP-
31	640	20.4	569	21 AAY83041	F-box protein hbet
32	640	20.4	569	21 AAY83250	Human cell signal
33	640	20.4	569	21 AAY44249	Human zF11 protein
34	637	20.3	569	22 AAB48298	Mouse ubiquitin li
35	635	20.2	569	21 AAB12812	F-box protein FWD1
36	635	20.2	569	21 AAY83254	WD-40 domain-contg
37	623	19.8	517	16 AAB85852	WD-40 domain-contg
38	618.5	19.7	779	16 AAB85854	F-box protein Cdo4
39	618.5	19.7	779	21 AAY83249	Human secreted pro
40	461	14.7	317	21 AAB63186	Human GTP-binding
41	461	14.7	334	22 AAB68529	OPDE 45 kba subuni
42	439	14.0	409	16 AAR70002	OPDE 45 kba subuni
43	439	14.0	410	16 AAR70005	WD-40 domain-contg
44	424.5	13.5	409	16 AAB85868	Gene 3 human secre
45	396	12.6	323	21 AAB63185	

#### ALIGNMENTS

#### RESULT 1

AAAY22466  
ID AAY22466 standard; Protein; 589 AA.  
XX  
AC AAY22466;  
XX  
DT 29-SEP-1999 (first entry)  
XX  
DE Human mammary sel-10 protein sequence.  
XX  
KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-4;  
KW mammary gland; therapy.  
XX  
OS Homo sapiens.  
XX  
PN W09932623-A1.  
XX  
PD 01-JUL-1999.  
XX  
PF 17-DEC-1998; 98WO-0526820.  
XX  
PR 19-DEC-1997; 97US-0068249.  
XX  
(PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Gurney ME, Li J, Pauley AM;  
XX  
DR WPI; 1999-458026/38.  
XX  
PT N-PSDB; AAX99702.  
XX  
PT New isolated human sel-10 polypeptides  
XX  
PS Claim 24; Page 60-63; 91pp; English.  
XX

CC This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human mammary sel-10 protein. The polypeptides  
CC can be used to alter presenilin function. Compounds which inhibit either  
CC the expression or the activity of the human sel-10 polypeptides may  
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and  
CC therefore may be useful for the prevention or treatment of Alzheimer's  
CC disease.

XX SQ Sequence 589 AA;

Query Match 100.0%; Score 3143; DB 20; Length 589;  
Best Local Similarity 100.0%; Pred. No. 1.1e-298;  
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKPGKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPEFCRRMRKLDHGSEV 60  
Db 1 MSKPGKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPEFCRRMRKLDHGSEV 60  
Qy 61 RSFSLGKPKCKVSEYTSSTGLVPCSATPTTFGDLRAANGQQRRTTSVQPTGLQEWL 120  
Db 61 RSFSLGKPKCKVSEYTSSTGLVPCSATPTTFGDLRAANGQQRRTTSVQPTGLQEWL 120  
Qy 121 KMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQFISLLPKELALYLSFLEP 180  
Db 121 KMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQFISLLPKELALYLSFLEP 180  
Qy 181 KDLQAAQTCRYWRILAEDNLLWRECKEEGIDEPLHTRKRVKPKGIHSPWKSAYIRQ 240  
Db 181 KDLQAAQTCRYWRILAEDNLLWRECKEEGIDEPLHTRKRVKPKGIHSPWKSAYIRQ 240  
Qy 241 HRIDTNWRRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLV 300  
Db 241 HRIDTNWRRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLV 300  
Qy 301 GHTGVSWSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360  
Db 301 GHTGVSWSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360  
Qy 361 RDATLRVWDIETGQCLHVMGHVAAVRCVQYDGRVRSVAGYDFVYKVMWDPETETCLHTLQ 420  
Db 361 RDATLRVWDIETGQCLHVMGHVAAVRCVQYDGRVRSVAGYDFVYKVMWDPETETCLHTLQ 420  
Qy 421 GHTNRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHSITSGMELKDNILVSGN 480  
Db 421 GHTNRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHSITSGMELKDNILVSGN 480  
Qy 481 ADSTVKIWDIKTGQCLQTPGPNKHQSVAVTCLQFNKNFVITSSDDGTGKLDLKTGEFIR 540  
Db 481 ADSTVKIWDIKTGQCLQTPGPNKHQSVAVTCLQFNKNFVITSSDDGTGKLDLKTGEFIR 540

RESULT 2

AA59198  
ID AAB59198 standard; protein; 589 AA.

XX AC AAB59198;

XX DT 23-MAR-2001 (first entry)

XX DE Human mammary sel-10-1 protein.

XX KW Sel-10; human; Alzheimer's disease; Abeta.

XX OS Homo sapiens.

XX PN WO200075328-A1.

XX XX 14-DEC-2000.

XX 23-MAY-2000; 2000WO-US09814.  
XX 09-JUN-1999; 99US-0328877.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Pauley AM, Gurney ME, Li J;  
XX WPI; 2001-102404/11.  
XX New human sel-10 polypeptides and their encoding polynucleotides,  
XX useful for raising antibodies for detecting sel-10 polypeptide  
XX expression and as drug targets in the treatment of Alzheimer's disease  
XX  
XX Claim 1; Page 79-82; 116pp; English.  
XX The present invention relates to human sel-10. The sel-10 proteins of  
XX the invention are useful for raising monoclonal or polyclonal  
XX antibodies useful in diagnostic assays for detecting sel-10  
XX polypeptide expression. The sel-10 polypeptides are also useful as drug  
XX targets for decreasing antibody levels in the treatment of Alzheimer's  
XX disease. It is also useful for identifying agents capable of  
XX altering the production level of Abeta. The polynucleotides are useful  
XX for developing assays for identifying agents capable of interfering  
XX with the biological pathways that lead to Alzheimer's disease.  
XX SQ Sequence 589 AA;

Query Match 100.0%; Score 3143; DB 22; Length 589;  
Best Local Similarity 100.0%; Pred. No. 1.1e-298;  
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKPGKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPEFCRRMRKLDHGSEV 60  
Db 1 MSKPGKPTLNHGLVPDLKSAKEPLPHQTVMKIFSIISIAOGLPEFCRRMRKLDHGSEV 60  
Qy 61 RSFSLGKPKCKVSEYTSSTGLVPCSATPTTFGDLRAANGQQRRTTSVQPTGLQEWL 120  
Db 61 RSFSLGKPKCKVSEYTSSTGLVPCSATPTTFGDLRAANGQQRRTTSVQPTGLQEWL 120  
Qy 121 KMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQFISLLPKELALYLSFLEP 180  
Db 121 KMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQRFQFISLLPKELALYLSFLEP 180  
Qy 181 KDLQAAQTCRYWRILAEDNLLWRECKEEGIDEPLHTRKRVKPKGIHSPWKSAYIRQ 240  
Db 181 KDLQAAQTCRYWRILAEDNLLWRECKEEGIDEPLHTRKRVKPKGIHSPWKSAYIRQ 240  
Qy 241 HRIDTNWRRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLV 300  
Db 241 HRIDTNWRRGELKSPKVLKGHDHVTITCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLV 300  
Qy 301 GHTGVSWSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360  
Db 301 GHTGVSWSQMRDNIISGSTDRTLKVNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360  
Qy 361 RDATLRVWDIETGQCLHVMGHVAAVRCVQYDGRVRSVAGYDFVYKVMWDPETETCLHTLQ 420  
Db 361 RDATLRVWDIETGQCLHVMGHVAAVRCVQYDGRVRSVAGYDFVYKVMWDPETETCLHTLQ 420  
Qy 421 GHTNRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHSITSGMELKDNILVSGN 480  
Db 421 GHTNRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHSITSGMELKDNILVSGN 480  
Qy 481 ADSTVKIWDIKTGQCLQTPGPNKHQSVAVTCLQFNKNFVITSSDDGTGKLDLKTGEFIR 540  
Db 481 ADSTVKIWDIKTGQCLQTPGPNKHQSVAVTCLQFNKNFVITSSDDGTGKLDLKTGEFIR 540  
Qy 541 NLVTLESQSGGVVWIRASNTKLVCAGSRNGTEETKLLVLDLDFVDMK 589  
Db 541 NLVTLESQSGGVVWIRASNTKLVCAGSRNGTEETKLLVLDLDFVDMK 589

Db 541 nlvtlesggsggvvrrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT 3

AAB01204  
ID AAB01204 standard; Protein; 589 AA.  
XX  
AC AAB01204;

XX 03-NOV-2000 (first entry)

XX Human GTPase associated protein-29.

XX Guanine nucleotide binding protein; GTP-binding protein; G-protein;  
KW GTPase; GTPase associated protein; GTPAP; cell proliferation;  
KW autoimmune; inflammatory; immune system disorder; cancer; AIDS;  
KW acquired immune deficiency syndrome; asthma; atherosclerosis;  
KW arthritis; systemic lupus erythematosus; psoriasis; human.

XX Homo sapiens.

XX WO200031263-A2.

XX 02-JUN-2000.

XX 23-NOV-1999; 99WO-US28013.

XX 23-NOV-1998; 98US-0109592.

PR 04-FEB-1999; 99US-0118610.

PR 06-APR-1999; 99US-0127990.

XX (INCY-) INCYTE PHARM INC.

XX Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR;

PI Yang J, Azimzai Y;

XX WPI; 2000-400073/34.

DR N-PSDB; AAA49199.

XX Human GTPase associated proteins, polynucleotides, and antibodies,

XX useful for diagnosing, preventing and treating various diseases such as

XX atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),

XX asthma, and autoimmune diseases -

XX Claim 1; Page 118-120; 144pp; English.

XX Human cDNA libraries from various tissues were screened for GTPase  
CC associated proteins (GTPAP). The present sequence is human  
CC GTPAP-29 protein. This sequence was derived from a cDNA library of  
CC brain tumour tissue. This protein is expressed in nervous,  
CC reproductive and gastrointestinal tissue. The GTPAP proteins may be  
CC used to define agonists and antagonists of GTPAP activity and to  
CC generate antibodies to GTPAP. This means the GTPAP proteins may be  
CC useful for treatment or prevention of diseases associated with GTPAP  
CC such as cell proliferation disorders, autoimmune disorders,  
CC inflammatory disorders, immune system disorders, cancer, AIDS, asthma,  
CC atherosclerosis, arthritis, systemic lupus erythematosus and psoriasis.

XX Sequence 589 AA;

Query Match 99.9%; Score 3140; DB 21; Length 589;  
Best Local Similarity 99.8%; Pred. No. 2.2e-298;  
Matches 589; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKPGKPTLNHGLVVDLSAKELPHQTMKIFSIISIAQGLPFCRRMRKRLDHGSEV 60

Db 1 mskpgkptlnhglvvpdlksakelphtqvmrifsisiiaqglpfcrrmrkrlhdgsev 60

QY 61 RSFSLGKKPKVSEYSTGLVPCSATPTTFGLRAANGQGQRRRITSVQPTGLQEWL 120

Db 61 rsfslgkpkvseysttglvpcsatpttfgdraangggqrrritsvqptglqewl 120

QY 121 KMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQFQDFISLLPKELALYVLSLEP 180  
Db 121 kmfqswwsgpekllaldelidsceptqvkhumqvilepqfdrdfisllpkelalyvlsfle 180  
QY 181 KDLQAAQTCTRYRILAEADNLLWRECKEKGIDEPHLIKRRKVIKPGFIHSPKWSAYIRQ 240  
Db 181 kdllqaatctcryrilaednllwreckeegidephlkrkrvikpgfihspkwsayirg 240  
QY 241 HRJDTNRRGELKSPKVLKGHDHVTCLQCGNRIIVSGDDNTLKVWSAVTGKCLRTLV 300  
Db 241 hridtnrrgelkspkvlkgdhvhtclqcgnrivsgsddnllkwsavtgkclrtlv 300  
QY 301 GHTGGVSSQMRDNIISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGS 360  
Db 301 ghgtggvssqmrdniisgstdrtlkvwnaetgecihtlyghtstvrclmhlhekrvvs 360  
QY 361 RDATLRWDIETGQCLHVLGMHVAARVCVOYDGRVVSAGYDFWVKWDDPETETCLHTLQ 420  
Db 361 rdatlrvwdietgqclhvlgmhvaavrcvdydgrvvvsagdfmwkvwdpetetclhtlq 420  
QY 421 GHTNRVYSIQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHSLTSGMELKDNILVSGN 480  
Db 421 ghtnrvisyiqfdgihvvsghltsirvwdvetgncihtltghsltsgmelkdnilvsgn 480  
QY 481 ADSTVKIWDIKTGQCLQTLQGPKNKHQSAVTCLOFNKFNVTSSDDGTVKLWDLTKTGFI 540  
Db 481 adstvkdwiktgtqcltqlggnkqhqsavtclfnkfnvtssddgtvklwdlktgfi 540  
QY 541 NLVTLESGGSGVWIRASNTKLVCAGSRNGTETFKLLVLDFDVMK 589  
Db 541 nlvtlesggsggvvrrirasntklvcavgsrngteetkllvldfdvdmk 589

RESULT 4

AAY22467

ID AAY22467 standard; Protein; 559 AA.

XX AAY22467;

XX 29-SEP-1999 (first entry)

XX Human mammary sel-10 protein sequence.

XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2;

XX mammary gland; therapy.

XX Homo sapiens.

XX WO932623-A1.

XX 01-JUL-1999.

XX 17-DEC-1998; 98WO-US26820.

XX 19-DEC-1997; 97US-0068243.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX Gurney ME, Li J, Pauley AM;

XX WPI; 1999-458026/38.

XX N-PSDB; AAX99702.

XX New isolated human sel-10 polypeptides

XX Claim 24; Page 63-66; 91pp; English.

XX This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human mammary sel-10 protein. The polypeptides  
CC can be used to alter presenilin function. Compounds which inhibit either  
CC the expression or the activity of the human sel-10 polypeptides may  
CC reverse the effects of mutations to presenilin-1 (PS-1) or PS-2, and  
CC therefore may be useful for the prevention or treatment of Alzheimer's

CC disease.  
XX XX  
SQ Sequence 559 AA;

Query Match 94.9%; Score 2983; DB 20; Length 559;  
Best Local Similarity 100.0%; Pred. No. 4.6e-283;  
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 MKIFSIISIAOGLPCRRMRKLDHGSEVRSFSLGKKPKCVSEYSTTGLVPCSATPTT 90  
Db 1 mkifsisiaaglpfcrrmrkldhgsevrslgkpkcvseysttstglvpcsatptt 60

Qy 91 FGDLEAANGCQQRRIITSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVVKHM 150  
Db 61 fgdleaaangcqqrritstsvqptglqewlkmfqsngpeklallalidelidscptqvkhh 120

Qy 151 MQVIEPQRFQFISLLPKELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRECKKEE 210  
Db 121 mqviepqrfqfislplkelaalyvlsflepkdlqaaqtcrywri-laednllwreкке 180

Qy 211 GIDELPHIKRRKVKIPGFIHSPWKSAYIRQHRIDTNWRGELKSPKVLKGHDHVVITCLQ 270  
Db 181 gidelphikrrkvikpgfihspwksayirqhridtnwrge-lkspkvlkgdhddhvitclq 240

Qy 271 FCGNRIVSGSDNTLKVMSAVTGKCLRTLVGHTGCVSWSSQMRDNIISGSTDRTLKVVNA 330  
Db 241 fcgnrivsgsdntlkvmsavtgkclrtlvghtgcvswssqmr-dniisgstdrtlkvwna 300

Qy 331 ETGECIHTLYGHTSTVRCMHLHKKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVCVQ 390  
Db 301 etgecihtlyghtstvr-cmhlhkkrvvsgsrdatlr-vwdietgqclhvlmghvaavrcvq 360

Qy 391 YDGRVVSAGYDFMVKVWDPETETCLHTLQGTNRVYSLQPDGIHVSGSLDTSIRVWDV 450  
Db 361 ydgrvvsgaydfmvkvwdpetetclhtlqgtnrvyqlqf-dgihvsgsl-dtsirvwdv 420

Qy 451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHQSAVT 510  
Db 421 etgncihtltghqsltsgmelkdnilvsgnadstvk-iwdikt-gqclqtlqgnpkhqsavt 480

Qy 511 CLOFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGVWVRIRASNTKLVCAVGS 570  
Db 481 clfnknfvitssddgtvklwldktgefirnlvtlesgsgv-vwvri-rasntklvcavgs 540

Qy 571 RNGTEETKLLVDFDVMK 589  
Db 541 rngteetkllvldfdvdmk 559

RESULT 5  
AAB59199  
ID AAB59199 standard; protein; 559 AA.  
XX  
AC AAB59199;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Human mammary sel-10-2 protein.  
XX  
KW Sel-10; human; Alzheimer's disease; Abeta.  
XX  
OS Homo sapiens.  
XX  
PN WO200075328-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 23-MAY-2000; 2000WO-US09814.  
XX  
PR 09-JUN-1999; 99US-0328877.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.

XX  
PI Pauley AM, Gurney ME, Li J;  
XX WPI; 2001-102404/11.  
XX  
PT New human sel-10 polypeptides and their encoding polynucleotides,  
PT useful for raising antibodies for detecting sel-10 polypeptide  
PT expression and as drug targets in the treatment of Alzheimer's disease  
PT  
XX  
PS Claim 1; Page 83-86; 116pp; English.  
XX  
CC The present invention relates to human sel-10. The sel-10 proteins of  
CC the invention are useful for raising monoclonal or polyclonal  
CC antibodies useful in diagnostic assays for detecting sel-10  
CC polypeptide expression. The sel-10 polypeptides are also useful as drug  
CC targets for decreasing antibody levels in the treatment of Alzheimer's  
CC disease. It is also useful for identifying agents capable of  
CC altering the production level of Abeta. The polynucleotides are useful  
CC for developing assays for identifying agents capable of interfering  
CC with the biological pathways that lead to Alzheimer's disease.  
XX  
SQ Sequence 559 AA;

Query Match 94.9%; Score 2983; DB 22; Length 559;  
Best Local Similarity 100.0%; Pred. No. 4.6e-283;  
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 MKIFSIISIAOGLPCRRMRKLDHGSEVRSFSLGKKPKCVSEYSTTGLVPCSATPTT 90  
Db 1 mkifsisiaaglpfcrrmrkldhgsevrslgkpkcvseysttstglvpcsatptt 60

Qy 91 FGDLEAANGCQQRRIITSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVVKHM 150  
Db 61 fgdleaaangcqqrritstsvqptglqewlkmfqsngpeklallalidelidscptqvkhh 120

Qy 151 MQVIEPQRFQFISLLPKELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRECKKEE 210  
Db 121 mqviepqrfqfislplkelaalyvlsflepkdlqaaqtcrywri-laednllwreкке 180

Qy 211 GIDELPHIKRRKVKIPGFIHSPWKSAYIRQHRIDTNWRGELKSPKVLKGHDHVVITCLQ 270  
Db 181 gidelphikrrkvikpgfihspwksayirqhridtnwrge-lkspkvlkgdhddhvitclq 240

Qy 271 FCGNRIVSGSDNTLKVMSAVTGKCLRTLVGHTGCVSWSSQMRDNIISGSTDRTLKVVNA 330  
Db 241 fcgnrivsgsdntlkvmsavtgkclrtlvghtgcvswssqmr-dniisgstdrtlkvwna 300

Qy 331 ETGECIHTLYGHTSTVRCMHLHKKRVVSGSRDATLRVWDIETGQCLHVLGMHVAARVCVQ 390  
Db 301 etgecihtlyghtstvr-cmhlhkkrvvsgsrdatlr-vwdietgqclhvlmghvaavrcvq 360

Qy 391 YDGRVVSAGYDFMVKVWDPETETCLHTLQGTNRVYSLQPDGIHVSGSLDTSIRVWDV 450  
Db 361 ydgrvvsgaydfmvkvwdpetetclhtlqgtnrvyqlqf-dgihvsgsl-dtsirvwdv 420

Qy 451 ETGNCIHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHQSAVT 510  
Db 421 etgncihtltghqsltsgmelkdnilvsgnadstvk-iwdikt-gqclqtlqgnpkhqsavt 480

Qy 511 CLOFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGVWVRIRASNTKLVCAVGS 570  
Db 481 clfnknfvitssddgtvklwldktgefirnlvtlesgsgv-vwvri-rasntklvcavgs 540

Qy 571 RNGTEETKLLVDFDVMK 589  
Db 541 rngteetkllvldfdvdmk 559

RESULT 6  
AAY22461  
ID AAY22461 standard; Protein; 627 AA.



XX AAY22461;  
XX 29-SEP-1999 (first entry)  
XX Human hippocampal sel-10 protein sequence.  
XX Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;  
KW therapy.  
XX Homo sapiens.  
XX OS  
XX PN W09932623-A1.  
XX PD 01-JUL-1999.  
XX PF 17-DEC-1998; 98WO-US26820.  
XX PR 19-DEC-1997; 97US-0068243.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX PA  
XX PI Gurney ME, Li J, Pauley AM;  
XX WPI; 1999-458026/38.  
XX DR N-PSDB; AAX99701.  
XX PT New isolated human sel-10 polypeptides  
XX PS Claim 24; Page 43-46; 9lpp; English.  
XX  
XX This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human hippocampal sel-10 protein. The  
CC polypeptides can be used to alter presenilin function. Compounds which  
CC inhibit either the expression or the activity of the human sel-10  
CC polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)  
CC or PS-2, and therefore may be useful for the prevention or treatment of  
CC Alzheimer's disease.  
XX  
XX Sequence 627 AA;  
XX  
Query Match 92.0%; Score 2890.5; DB 20; Length 627;  
Best Local Similarity 89.9%; Pred. No. 6.3e-274;  
Matches 551; Conservative 11; Mismatches 10; Indels 41; Gaps 4;  
QY 13 LVPVDLSAKEPLPHQVTMKIFSIIIA-----OGLPFCRR----- 48  
DB 20 lllpvlpn-----lpflclsmstlesvtyipekgl-ycqrlpsrthggtesikgkntn 74  
QY 49 -----RMKRKLHDHSGSEVSFSLGKPKCKYSEYTSITGLVPCSATPTTFGLRA 96  
DB 75 mgfygtlkmifykmkrklidhgsevsfslgkpkckyseysttstglvpcsatpttfglra 134  
QY 97 ANGGQORRRITSVQPTGLQEWLKMFSQSGPEKLIALDELIDSCPTOVKHMVIEP 156  
DB 135 anggqorrritvqptglqewlkmfsgswgpeklialdelidscptovkhumviep 194  
QY 157 QFQRFISLLPKELALVLSFLEPKDLLQAAQTCRYWRILAEONLWRECKEEDIEPL 216  
DB 195 qfqrdfisllpkelalvlsflepdkllgaqtcrywriilaednllwrekkeegidepl 254  
QY 217 HIKRRVKIKGFTHSPPKSAIIRQHRTDTHWRGELKSPKVLKGHDHVTICLQFCGNRI 276  
DB 255 hikrrvkiqpfthspkksayirqhridtnwrrgelkspkvlkgdhvhticlqfcgnri 314  
QY 277 VSGSDNTLKVWSAVTGKCLRTLVGHTGGVWSSOMRNIILISGTDRTLKVWNAETGECI 336  
DB 315 vsgsdntlkvwsavtgkclrtlvghtggvwsqmqrdnliisgtdrtlkvwnaetgeci 374  
QY 337 HTLYGHTSTVRCMHLHKEKRVVSGSRDATLRWBDIETGCQLHLVLMGHVAARVCVQDGRV 396  
DB 375 htlyghtstvrchlhkekrvsvgsrdatlrwbdietgcqlhvlmghvaavrcvqydgrrv 434

QY 397 VSGAYDFWVKVWDPTETETCLHTLQGHTRNRYVSLOFDGHHVYVSGSIDTSIRYWDVETGNCI 456  
DB 435 vsgeydfmvkvwdptetctclhtlqghcnrnyvslofdghvsvgsidsirvwdvetgnci 494  
QY 457 HTLTGHQSLSGMEKDKNIIYVSGNADSTVKIWDIKTQCCLQTLOGPNKHQSASVTCLOFNK 516  
DB 495 htltghqsltsgmekdkniiyvgnadstvkliwdiktgcclqtlogpnkhqsavtclqfnk 554  
QY 517 NFVITSSDDGTGTVKLWDLKTGEFIRNLVLTESGGGVVWRIRASNTKLVCAGSRNGTTEE 576  
DB 555 nfvitssddgtgvtvklwldktgefirnlvtlesggvgvwrirasntklvcavsrngttee 614  
QY 577 TKLLVLDLDFDVMK 589  
DB 615 tkllvldldfvdvmk 627  
RESULT 7  
AAB59193  
ID AAB59193 standard; protein; 627 AA.  
XX AC  
XX AAB59193;  
XX 23-MAR-2001 (first entry)  
XX Human hippocampal sel-10-1 protein.  
XX Sel-10; human; Alzheimer's disease; Abeta.  
XX Homo sapiens.  
XX W0200075328-A1.  
XX PD 14-DEC-2000.  
XX PF 23-MAY-2000; 2000WO-US09814.  
XX PR 09-JUN-1999; 99US-0328877.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Pauley AM, Gurney ME, Li J;  
XX WPI; 2001-102404/11.  
XX New human sel-10 polypeptides and their encoding polynucleotides,  
XX useful for raising antibodies for detecting sel-10 polypeptide  
XX expression and as drug targets in the treatment of Alzheimer's disease  
XX  
XX Claim 1; Page 60-64; 116pp; English.  
XX The present invention relates to human sel-10. The sel-10 proteins of  
XX the invention are useful for raising monoclonal or polyclonal  
XX antibodies useful in diagnostic assays for detecting sel-10  
XX polypeptide expression. The sel-10 polypeptides are also useful as drug  
XX targets for decreasing antibody levels in the treatment of Alzheimer's  
XX disease. It is also useful for identifying agents capable of  
XX altering the production level of Abeta. The polynucleotides are useful  
XX for developing assays for identifying agents capable of interfering  
XX with the biological pathways that lead to Alzheimer's disease.  
XX  
XX Sequence 627 AA;  
XX

Query Match 92.0%; Score 2890.5; DB 22; Length 627;  
Best Local Similarity 89.9%; Pred. No. 6.3e-274;  
Matches 551; Conservative 11; Mismatches 10; Indels 41; Gaps 4;  
QY 13 LVPVDLSAKEPLPHQVTMKIFSIIIA-----OGLPFCRR----- 48  
DB 20 lllpvlpn-----lpflclsmstlesvtyipekgl-ycqrlpsrthggtesikgkntn 74

```
QY 49 -----RMKRKLDHGSEVRSFSLGKKPKCVSEYTTSTGLVPCSATPTTFGDLRA 96
Db :|||||
75 mgyfytllkmifkymkrkldhgsevrslgkpkckvseyststglvpcsatpttfgdrla 134
QY 97 ANGQGOORRRITSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVVKHMMQVIEP 156
Db :|||||
135 angqgqrrritsvqptglqewlkmfqsqgpekllalidelidsceptqvkhmmqviep 194
QY 157 QPQRFISLLPELALYVLSFLEPKDLQAQTCRYWRILAEADNLLWRKCKEEGIDEPL 216
Db :|||||
195 qfqrdfisllpelalylvlsflepkdlqaaqtcrywri-laednllwrekc-keegidepl 254
QY 217 HIKRRKVIKPGFIHSPWKSAYIROHRIDTNWRGELKSPKVLKGHDDHVTICLQFCGNRI 276
Db :|||||
255 hkrkrrvikipgfihs-ppwksayir-qhrdntwnrrgelkspkvlkgdh-dhvticlcfcgnri 314
QY 277 VSGSDNTLLKVMASVATGKCLRTLVLGHTGGVSSQMRDNIISGSTDRTLKVNAAETGECI 336
Db :|||||
315 vsgsdntllkvm-asvatgkclrtlvghtggvssqmr-dniisgst-drtlkv-naetgeci 374
QY 337 HTLYGHTSTVRCMHLHKKRVVSGSRDATLRVWDIETGQCLHVLGMGHVAAVRCVQYDGRV 396
Db :|||||
375 htlyghtstvr-cmhlh-ekrvvsgs-datlr-vwdietg-qclhvlgmgh-vaavrcv-qydgrrv 434
QY 397 VSGAYDFWVKVWDPEPETETCLHTLQGHTRNRYVSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
Db :|||||
435 vsgaydfwmkv-wdpet-etclhtlqghtrnrvysl-qfdgihvvs-gsl-dtsirv-wdvetgnci 494
QY 457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHOSAVTCLQFNK 516
Db :|||||
495 htltghqslts-gmelk-dnilvsgnadstvk-iwdikt-gqclqlt-lqgnpkh-osavt-clqfnk 554
QY 517 NFVITSSDDGTVKLWDLKTGEFIRNLVLTLESGGSGGVVWRIRASNTKLVCVAGSRNGTEE 576
Db :|||||
555 nfvitssddgtvkl-wdlkt-gefir-nlvltlesgsggv-vwrir-asntklvc-agsrngtee 614
QY 577 TKLLVLDFFDVMK 589
Db :|||||
615 tkllvldf-dvmk 627

RESULT 8
AAY22471
ID AAY22471 standard; Protein; 666 AA.
XX AC AAY22471;
XX DT 29-SEP-1999 (first entry)
XX DE Human C-term mychis tagged sel-10 protein sequence.
XX KW Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;
XX KW therapy; mammary gland.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN W09932623-A1.
XX PD 01-JUL-1999.
XX PF 17-DEC-1998; 98WO-US26820.
XX PR 19-DEC-1997; 97US-0068243.
XX PA (PHAA ) PHARMACIA & UPJOHN CO.
XX PI Gurney ME, Li J, Pauley AM;
XX DR WPI: 1999-458026/38.
XX DR N-PSDB; AAX99716.
```

```
PT New isolated human sel-10 polypeptides
XX Example 3; Page 85-88; 91pp; English.
PS
XX This sequence is an epitope-tagged version of a human sel-10 protein
CC of the invention. The human sel-10 proteins of the invention are isolated
CC from hippocampus and mammary gland. The polypeptides can be used to
CC alter presenilin function. Compounds which inhibit either the expression
CC or the activity of the human sel-10 polypeptides may reverse the effects
CC of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
CC for the prevention or treatment of Alzheimer's disease.
XX Sequence 666 AA:

Query Match 92.0%; Score 2890.5; DB 20; Length 666;
Best Local Similarity 89.9%; Pred. No. 7e-274;
Matches 551; Conservative 11; Mismatches 10; Indels 41; Gaps 4;

QY 13 LYPVDLKSNAKEPLPHQTYVMKIFSIISIIA-----QGLPFCRR----- 48
Db :|||||
20 lipvlpn-----lpftclsmstlesvtlpekgl-ycqrlpssrthgteslkgkntn 74
QY 49 -----RMKRKLDHGSEVRSFSLGKKPKCVSEYTTSTGLVPCSATPTTFGDLRA 96
Db :|||||
75 mgyfytllkmifkymkrkldhgsevrslgkpkckvseyststglvpcsatpttfgdrla 134
QY 97 ANGQGOORRRITSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVVKHMMQVIEP 156
Db :|||||
135 angqgqrrritsvqptglqewlkmfqsqgpekllalidelidsceptqvkhmmqviep 194
QY 157 QPQRFISLLPELALYVLSFLEPKDLQAQTCRYWRILAEADNLLWRKCKEEGIDEPL 216
Db :|||||
195 qfqrdfisllpelalylvlsflepkdlqaaqtcrywri-laednllwrekc-keegidepl 254
QY 217 HIKRRKVIKPGFIHSPWKSAYIROHRIDTNWRGELKSPKVLKGHDDHVTICLQFCGNRI 276
Db :|||||
255 hkrkrrvikipgfihs-ppwksayir-qhrdntwnrrgelkspkvlkgdh-dhvticlcfcgnri 314
QY 277 VSGSDNTLLKVMASVATGKCLRTLVLGHTGGVSSQMRDNIISGSTDRTLKVNAAETGECI 336
Db :|||||
315 vsgsdntllkvm-asvatgkclrtlvghtggvssqmr-dniisgst-drtlkv-naetgeci 374
QY 337 HTLYGHTSTVRCMHLHKKRVVSGSRDATLRVWDIETGQCLHVLGMGHVAAVRCVQYDGRV 396
Db :|||||
375 htlyghtstvr-cmhlh-ekrvvsgs-datlr-vwdietg-qclhvlgmgh-vaavrcv-qydgrrv 434
QY 397 VSGAYDFWVKVWDPEPETETCLHTLQGHTRNRYVSLQFDGIHVVSGSLDTSIRVWDVETGNCI 456
Db :|||||
435 vsgaydfwmkv-wdpet-etclhtlqghtrnrvysl-qfdgihvvs-gsl-dtsirv-wdvetgnci 494
QY 457 HTLTGHQSLTSGMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGNPKHOSAVTCLQFNK 516
Db :|||||
495 htltghqslts-gmelk-dnilvsgnadstvk-iwdikt-gqclqlt-lqgnpkh-osavt-clqfnk 554
QY 517 NFVITSSDDGTVKLWDLKTGEFIRNLVLTLESGGSGGVVWRIRASNTKLVCVAGSRNGTEE 576
Db :|||||
555 nfvitssddgtvkl-wdlkt-gefir-nlvltlesgsggv-vwrir-asntklvc-agsrngtee 614
QY 577 TKLLVLDFFDVMK 589
Db :|||||
615 tkllvldf-dvmk 627

RESULT 9
AAB59203
ID AAB59203 standard; protein; 666 AA.
XX AC AAB59203;
XX DT 23-MAR-2001 (first entry)
XX DE C-terminal mychis tagged sel-10 protein.
```

D	b	435	vsgaydfmvkwbdpetetclhtlqgthnrvyslqfdgihvsgsldtsirvwdvetgnci	494
Q	y	457	HTLTGHOSLTSGMELKDNILVSGNADSTVKINDIKTGQCLOTLQGNKHQSAVTCLQFNK	516
D	b	495	htlctghqslctsgmelkdnillvsgnadstvkldwiktggclgtlqgnkhqgsavtclqfnk	554
Q	y	517	NFVITSSDDGTVKLMDLKTGEFIRNLVTLESGGGVVWRIRASNTKLVCAVGSRRNGTEE	576
D	b	555	nfvitssddgtvklwdlktgefirnlvtlesggsggvvwrirastklvcavgsrrngtee	614
Q	y	577	TKLLVLDFDMDK	589
D	b	615	tkllvldfdvdmk	627
RESULT 10				
A	A	Y22470		
I	D	AAY22470	standard; Protein; 669 AA.	
X	X	AC	AAY22470;	
X	X	DT	29-SEP-1999 (first entry)	
X	X	DE	Human Cterm V5 his tagged sel-10 protein sequence.	
X	X	KW	Sel-10; human; presenilin; Alzheimer's disease; PS-1; PS-2; hippocampus;	
X	X	KW	therapy; mammary gland.	
X	X	OS	Homo sapiens.	
X	X	OS	Synthetic.	
X	X	PN	W09932623-A1.	
X	X	PD	01-JUL-1999.	
X	X	PF	17-DEC-1998; 98WO-US26820.	
X	X	YY		

```

XX      (PHAA ) PHARMACIA & UPJOHN CO.
XX PA
XX PI
XX PT Gurney ME, Li J, Pauley AM;
XX DR WPI; 1999-458026/38.
XX DR N-PSDB; AAX99715.
XX PT New isolated human sel-10 polypeptides
XX PS Example 3; Page 79-83; 91pp; English.
XX CC This sequence is an epitope-tagged version of a human sel-10 protein
CC of the invention. The human sel-10 proteins of the invention are isolated
CC from hippocampus and mammary gland. The polypeptides can be used to
CC alter presenilin function. Compounds which inhibit either the expression
CC or the activity of the human sel-10 polypeptides may reverse the effects
CC of mutations to presenilin-1 (PS-1) or PS-2, and therefore may be useful
CC for the prevention or treatment of Alzheimer's disease.
XX SQ Sequence 669 AA;

Query Match          92.0%; Score 2890.5; DB 20; Length 669;
Best Local Similarity 89.9%; Pred.No. 7e-274;
Matches 551; Conservative 11; Mismatches 10; Indels 41; Gaps 4

QY    13 LVPVDLKSAREPLHPTVMKIFSIITIA----QGLPFCCR----- 48
       ||| : || | : : : : || :||
Db     20 llpyllon---lpftclsmstlesvtylpekgl-ycarlpssrthagtslkqknten 74

QY    49 -----RMKKRLDHGSEVRSFSLGGKPKCKVEYSTTTLGVPCSATPTTFGDLRA 96
       :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     75 mgfygtlkmifymkr-kldhgsevrfsfslgkkpkckveytsttlgvpcsatpttfgdla 134

```



XX Homo sapiens.  
OS  
XX  
PN W0932623-A1.  
XX  
XX  
PD 01-JUL-1999.  
XX  
PF 17-DEC-1998; 98WO-US26820.  
XX  
PR 19-DEC-1997; 97US-0068243.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Gurney ME, Li J, Pauley AM;  
XX  
DR WPI; 1999-458026/38.  
DR N-PSDB; AAX99701.  
XX  
XX New isolated human sel-10 polypeptides  
PT  
XX  
PS Claim 24; Page 47-50; 91pp; English.  
XX  
XX This sequence represents a human sel-10 protein of the invention. This  
CC sequence is specifically a human hippocampal sel-10 protein. The  
CC polypeptides can be used to alter presenilin function. Compounds which  
CC inhibit either the expression or the activity of the human sel-10  
CC polypeptides may reverse the effects of mutations to presenilin-1 (PS-1)  
CC or PS-2, and therefore may be useful for the prevention or treatment of  
CC Alzheimer's disease.  
XX  
XX Sequence 592 AA;  
SQ

Query Match 92.0%; Score 2890; DB 20; Length 592;  
Best Local Similarity 93.4%; Pred. No. 6.5e-274;  
Matches 549; Conservative 3; Mismatches 10; Indels 26; Gaps 2;  
QY 2 SRPGKPTLNHGLVPVDLKSACEPLPHQTVMKIFSIISIAAGLPFCRRMRKRLDHGSEVR 61  
DB 31 slkgkntnmgf-----ygtlkmif-----yknkrklidhgsevr 64

QY 62 SPSLGKKPKVSEYTSSTGLVPCSATPTTFGLRAANGOGQRRRTITSVQPTGLQEWLK 121  
DB 65 sfslgkpkckvseystatglvpcsatpttfgdlraangggqrrritsvpptglqewlk 124

QY 122 MFQSGSGPEKLLALDELIDSCPTQVKHMQVTEPQFQDFISLLPKELALYVLSFLEPK 181  
DB 125 mfgswsgpeklaladelidsceptqvkmmqvleqpqfdrfislipkelalylvlsflepk 184

QY 182 DILQAQTCRYWRILAEDNLLWREKCEGIDEPHKKRRKVIKPGFIHSPWKSAYIROH 241  
DB 185 dillqaqtcrywri-laednllwrekeegidephikrrkvikpgfi-hspwksayirqh 244

QY 242 RIDTNRRRELKSPVKLGHDHVTICLQFCGNRIYVSGDDNTLKWSAVTGKCLRTLVG 301  
DB 245 ridtnwrrgelkspvklghdhvhtclqfcgnriyvsdndtlkwsvavgkclrtlv 304

QY 302 HTGGVWSSQMRDNIISGSTDRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSR 361  
DB 305 htggvwsqmrdniisgstdrtlkwnnaetgecihtlyghtstvr-cmhlhekrvvsgr 364

QY 362 DATLRWIDLETGQCILHVLHMGHVAARVCQYDGRVVVSGAYDFMVKVWDPETETCLHTLQ 421  
DB 365 datlrwldetgqclhvlmgghvaarvcydgrrvvvsaydfm-vkwvdpetetc-lhtlq 424

QY 422 HTNRYVSLQFDGIHVYVSGSLDTSIRWVDVETGNCIHTLTGHQSLTSGMELKNILVSGNA 481  
DB 425 hcnryvslqfdgi-hvyvsgsl-dtsirwvdetgncihtltghqsl-tsgmelk-nilvsgna 484

QY 482 DSTVKIWDIKTCQCLQTLQGPKNHQSVAVTCLOFNKNFVITSSDDGTIVKLWDLKTGEFIRN 541  
DB 485 dstvkiwdiktqclqtlqgpnkhqsavtclofnknfnvittssddgtiv-klwdlktgefirn 544

QY 542 LVTLESGGGVVWRIRASNTKLVCAGSRNGTEETKLLVLDFDVMK 589  
DB 545 lvtlesggsggvvwrir-asntklvcavsrngteetkllvldfdvdmk 592

RESULT 13  
AAB59194  
ID AAB59194 standard; protein; 592 AA.  
XX  
AC AAB59194;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Human hippocampal sel-10-2 protein.  
XX  
KW Sel-10; human; Alzheimer's disease; Abeta.  
XX  
OS Homo sapiens.  
XX  
PN W0200075328-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 23-MAY-2000; 2000WO-US09814.  
XX  
PR 09-JUN-1999; 99US-0328877.  
XX  
PA (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
PI Pauley AM, Gurney ME, Li J;  
XX  
DR WPI; 2001-102404/11.  
XX  
PT New human sel-10 polypeptides and their encoding polynucleotides,  
PT useful for raising antibodies for detecting sel-10 polypeptide  
PT expression and as drug targets in the treatment of Alzheimer's disease  
PT  
XX  
XX Claim 1; Page 64-68; 116pp; English.  
XX  
CC The present invention relates to human sel-10. The sel-10 proteins of  
CC the invention are useful for raising monoclonal or polyclonal  
CC antibodies useful in diagnostic assays for detecting sel-10  
CC polypeptide expression. The sel-10 polypeptides are also useful as drug  
CC targets for decreasing antibody levels in the treatment of Alzheimer's  
CC disease. It is also useful for identifying agents capable of  
CC altering the production level of Abeta. The polynucleotides are useful  
CC for developing assays for identifying agents capable of interfering  
CC with the biological pathways that lead to Alzheimer's disease.  
XX  
SQ Sequence 592 AA;  
QY 2 SRPGKPTLNHGLVPVDLKSACEPLPHQTVMKIFSIISIAAGLPFCRRMRKRLDHGSEVR 61  
DB 31 slkgkntnmgf-----ygtlkmif-----yknkrklidhgsevr 64

QY 62 SPSLGKKPKVSEYTSSTGLVPCSATPTTFGLRAANGOGQRRRTITSVQPTGLQEWLK 121  
DB 65 sfslgkpkckvseystatglvpcsatpttfgdlraangggqrrritsvpptglqewlk 124

QY 122 MFQSGSGPEKLLALDELIDSCPTQVKHMQVTEPQFQDFISLLPKELALYVLSFLEPK 181  
DB 125 mfgswsgpeklaladelidsceptqvkmmqvleqpqfdrfislipkelalylvlsflepk 184

QY 182 DILQAQTCRYWRILAEDNLLWREKCEGIDEPHKKRRKVIKPGFIHSPWKSAYIROH 241  
DB 185 dillqaqtcrywri-laednllwrekeegidephikrrkvikpgfi-hspwksayirqh 244



PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Claim 1; Page 72-75; 116pp; English.

The present invention relates to human sel-10. The sel-10 proteins of the invention are useful for raising monoclonal or polyclonal antibodies useful in diagnostic assays for detecting sel-10 polypeptide expression. The sel-10 polypeptides are also useful as drug targets for decreasing antibody levels in the treatment of Alzheimer's disease. It is also useful for identifying agents capable of altering the production level of Abeta. The polynucleotides are useful for developing assays for identifying agents capable of interfering with the biological pathways that lead to Alzheimer's disease.

Sequence 545 AA;

Query Match 91.9%; Score 2889; DB 22; Length 545;

Best Local Similarity 99.8%; Pred. No. 7.le-274;

Matches 540; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 49 RMKRKLDHGSEVRSFSLGKKPKCVSEYTTGLVPGSATPTTFGDLRAANGCQQRRIIT 108

Db :||||| 5 kmrkldhgsevrfsfslgkpkcvseysttglvpcsatpttfgdlaangqgqrrrit 64

QY 109 SVQPPPTGLQEWLKMFSWGSQPEKLLALDELIDSCPTQVKHMQVIEPOFDFISLLPK 168

Db :||||| 65 svqptglqewlkmfswgsqpeklaldelidsceptqvkmmqvlepfqrdfisllpk 124

QY 169 ELALYVLSFLEPKDLLOAQTCHRYRILAEADNLLWRECKEKEGIDEPHLIKRKVKIKPGF 228

Db :||||| 125 elalyvlsflepkdlloaaqtcrywilaednllwreckeegideplhikrrkvikpgf 184

QY 229 IHSPWKSAYITROHRIDTNRRGELKSPKVLKGHDDHVITCLQPCGNRIYSGSDNTLKVW 288

Db :||||| 185 ihspwksayirhridtnrrrgelkspkvlkgdhdhvitclqfcgnriyvsdndtlkvw 244

QY 289 SAVTGKCLRTLVGHTGVWSSQMRDNIISGSDRTLKVNAETGECIHTLYGHTSTVRC 348

Db :||||| 245 savtgkclrtlvghtgvwssqmrndniisgstdrtlkvnaetgecihtlyghtstvrc 304

QY 349 MHLHRRVYSGSRDATLRVWDIETGQCLHVLGMHVAARVCQYDGRVRVSGAYDFMVKVW 408

Db :||||| 305 mhlhrrvysgsrdatlrwdietgqclhvlgmhvaavrcvdygrvrsvgaydfmfvkw 364

QY 409 DPETETCLHTLQGHTRVYSLQFDGIHVYSGSLDTSIRVWDVETGNCIHTLGHQSLTSG 468

Db :||||| 365 dpetetclhtlqghtrvyslqfdgihvysgsldtsirvwdvetgncihtlghqslts 424

QY 469 MELKDNILVSGNADSTVKIWDIKTGCLQTLQCPNKHQSAVTCLOFNKFNVTSSDDGTV 528

Db :||||| 425 melkdnilvsgnadstvkiwdiktgclqtlqcpnkhgsavtclofnkfnvtsddgtv 484

QY 529 KLWDLKTGFIRNLVTLESGGGVVWRIRASNTKLVCVSGSRNGTEETKLLVLDVDM 588

Db :||||| 485 klwldktgfirnlvtlesgggvvwrirasntklvcavsgsrngteetkllvldfvd 544

QY 589 K 589

Db : 545 k 545

Search completed: September 25, 2001, 14:41:10  
Job time: 165 sec

**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	1238	39.4	587	3	US-08-899-578-2	Sequence 2, Appli	
2	671	21.3	209	3	US-08-899-578-6	Sequence 6, Appli	
3	623	19.8	517	1	US-08-190-802A-30	Sequence 30, Appl	
4	618.5	19.7	779	1	US-08-190-802A-32	Sequence 32, Appl	
5	439	14.0	409	2	US-08-283-917-3	Sequence 3, Appli	
6	439	14.0	409	2	US-08-961-716-3	Sequence 3, Appli	
7	439	14.0	410	2	US-08-283-917-9	Sequence 9, Appli	
8	439	14.0	410	2	US-08-961-716-9	Sequence 9, Appli	
9	424.5	13.5	409	1	US-08-190-802A-51	Sequence 51, Appl	
10	410	13.0	212	3	US-08-899-578-7	Sequence 7, Appli	
11	376.5	12.0	422	1	US-08-190-802A-52	Sequence 52, Appl	
12	364.5	11.6	514	1	US-08-190-802A-56	Sequence 66, Appl	
13	332	10.6	704	1	US-08-188-582-5	Sequence 5, Appli	
14	332	10.6	704	1	US-08-646-715-5	Sequence 5, Appli	
15	326	10.4	704	1	US-08-190-802A-62	Sequence 62, Appl	
16	326	10.4	704	2	US-08-308-818-3	Sequence 3, Appli	
17	319.5	10.2	713	1	US-08-190-802A-63	Sequence 63, Appl	
18	317.5	10.1	704	1	US-08-188-582-18	Sequence 18, Appl	
19	317.5	10.1	704	1	US-08-646-715-18	Sequence 18, Appl	
20	311.5	9.9	375	4	US-09-063-743-1	Sequence 1, Appli	
21	306	9.7	798	1	US-08-190-802A-64	Sequence 1, Appli	
22	306	9.7	798	1	US-08-190-802A-68	Sequence 68, Appl	
23	306	9.7	798	2	US-08-308-818-2	Sequence 2, Appli	
24	288	9.2	318	1	US-08-190-802A-33	Sequence 33, Appl	
25	287	9.1	906	1	US-08-190-802A-31	Sequence 31, Appl	
26	280.5	8.9	340	1	US-08-190-802A-38	Sequence 38, Appl	
27	276	8.8	340	4	US-09-180-783-2	Sequence 2, Appli	

```
Db 77 -----VEHLIAFKDLSSAEQMDAFRLQLQESNMNINQLRAITIEPHFQDRDLSCLPVEL 131
Qy 171 ALYVLSFLEPKDQLQAQTCRYWRILAEADNLLWREKCKEE-----210
Db 132 GKKILHNHTGYDLLKVAQVSKNMKLISEIDIKWISLGSVEEFKHPDPDTRVTGAWQGTAI 191
Qy 211 --GIDEPLHI-----KRRKVIKPG--FIHSPWKSAYIRQHRTDTNRRRGELKSPKVL 258
Db 192 AGVTPIPHIQPCDNLNVRFLKQLKQFGDIFERAADKSYLRADKIEKNWNPIMGSAYL 251
Qy 259 KGHDDHVITCLQFCGNRIYSGDDNTLKWASAVTGKCLRTLVGHTGGVWSSQMRD--NII 316
Db 252 RHEDHVITCQIHDVLTGSDONTLKWICDKGEVMTLVGHTGGVWTSQCGRYI 311
Qy 317 IGSGLDRTLVKVNNAETGECIHTLVGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCL 376
Db 312 VSGSDTRVTKVMSTVDGSLHTLOQHTSTVRCMAMAGSILVTGSRDRTLVRVWDESGRHL 371
Qy 377 HVLGCHVAARVCQVQDGRVYSGAYDFMVKVMDPETETCLHTLOQHTNRVYSLOQFDGIH- 435
Db 372 ATLGHGHAARVCQVQDGTTVVSGGYDFTVKIWNHATGRCIRTLGHNNRVYSLFESERS 431
Qy 436 --VSGSLDTSIRVWDV---ETGNCIHTLTGHOSLTSGMELKDNILVSGNADSTVKIWDIK 491
Db 432 IVCSGLDTSIRVWDFTRPEQGEVALLQHTSLTSGMQLRGNILVSGNADSHVRVWDIH 491
Qy 492 TGQCLOTLQGNPKHQSAYTCLQ-FPNKNFVITSSDDGTVKLWDLKTGEFIRNLVLTLESGS 550
Db 492 EGTCTVHMLSG---HRSATISLQWFGNRNVATSSDDGTVKLWDIERGALLIRDLVLTLDSSGN 548
Qy 551 GGVWRIRASNTKLVCAVSGSRNGTEETKLLVLDFFD 585
Db 549 GGCWIRLCTSTMLACAVSGSRNNTETKVLDDFD 583
```

## RESULT 2

```
US-08-899-578-6
; Sequence 6, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-899-578-6

Query Match 21.3%; Score 671; DB 3; Length 209;
Best Local Similarity 52.2%; Pred. No. 8.3e-60;
Matches 145; Conservative 24; Mismatches 35; Indels 74; Gaps 9;

Qy 260 GHDDHVITCLQFCGNRIYSGDDNTLKWASAVTGKCLRTLVGHTGGVWSSQMRD--NII 317
Db 1 GHEDHVITCQIHDVLTGSDONTLKW-----CGHTGGVWTSQCGRYIV 49
Qy 318 SGSDTRTLKVNNAETGECIHTLVGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLH 377
Db 50 SGSDTRTVKWS-----GHTSTVRCMAMAGSILVTGSRDRTLVRVWD-----90
Qy 378 VLMGHVAARVCQVQDGRVYSGAYDFMVKVMDPETETCLHTLOQHTNRVYSLOQFDGIH-- 435
Db 91 ---GHAAARVCQVQDGTTVVSGGYDFTVKIWN-----GHNRVYSLFESERSI 136
Qy 436 VVSGSLDTSIRVWDVETGNCIHTLTGHOSLTSGMELKDNILVSGNADSTVKIWDIKTQOC 495
Db 137 VCSGLDTSIRVWD-----GHTSLTSGMQLRGNILVSGNADSHVRVWD-----179
Qy 496 LQTLQGNPKHQSAYTCLQ-FPNKNFVITSSDDGTVKLWD 532
Db 180 -----GHSATISLQWFGNRNVATSSDDGTVKLWD 209

RESULT 3
US-08-190-802A-30
; Sequence 30, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
```

US-08-190-802A-30

Query Match 19.8%; Score 623; DB 1; Length 517;  
Best Local Similarity 31.1%; Pred. No. 2.6e-54;  
Matches 161; Conservative 89; Mismatches 214; Indels 54; Gaps 13;

QY 60 VRSFSLGKPKCKSEYTSITGLVPCSATPT---TFGDLRAANGQGOQ---RRRITSVQPP 113  
DB 1 MEGFSCSLQPTASERDCNRDEPRKITEKNTLQTKLANGTSMIVPKQRKLSANYE 60  
QY 114 TGLQWLKMFQSNSEKLLALDELIDSEPTQVKHMMQVIEPQFORDFISLLP-----KE 169  
DB 61 KEKELCVKYEQNSCEDQVEFVHLRLSRMCHYQGHGINTYKPLQRFDTALPARGLDH 120  
QY 170 LALVLSFLEPKDLQAQOTCRVWRILAEENLLWRECKE-----EGIDPELHKRR 221  
DB 121 IAEENLLWRECKE-----EGIDPELHKRR 221  
QY 222 K-----VIKPGFIHSPKSAI-----IRQ---HRIDTNMRGELKSPKV-LKGHDHVI 266  
DB 176 GWGQYLFKNKPPDKPPNSFYRALYPKLIQDIETIESNRCGRHSIORHCHSETSKGV 235  
QY 267 TGLQFNGNRIVSGSDNTLKVSAVTKGLRLTLVHTGGVWSSQMRDNIISGSDRTLK 326  
DB 236 YCLQYDDQKIVSGLRDNTIKINDKNTLECKRVLMTGHTGSLVCLQYDERVITGS-DSTVR 294  
QY 327 VNAETGECIHTLYGHTSTVRCHLHEKRVVSGSRDATLRVWDIETGQCL---HVLMGHV 383  
DB 295 VMDVTGEMTLTHHCEAVLHFRNNGMMVTCSKRSIAVMDMASATDITLRRVLVGRH 354  
QY 384 AAVRCVQYDGRVVSAYDPMVWVDPETETCLHTLQHTNRVYSLQFDGTHVVSGLDT 443  
DB 355 AAVNVVDFDDKYIVSAGDRTIKVWTSTCEVFTLNGHKGACIQLQYDRDLVVSQSDN 414  
QY 444 SIRVDVETGNCIHTLTHGSLFSGMELKNDNLVSGNADSTVKIMDKTG-----Q 494  
DB 415 TIRLWDIEGACLRVLEHGEELVRCIRFONKRVISGAYDKIKVMDLVAALDPRAPAGTL 474  
QY 495 CLTQLOGPNKHOSAVTCLQFNKFNVTSSDDGTGVKLWD 532  
DB 475 CLRTLV---EHSGRVPRLOFDEQIVSSSHDDTILWD 509

RESULT 4  
US-08-190-802A-32  
Sequence 32, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron, Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 779 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15  
US-08-190-802A-32

Query Match 19.7%; Score 618.5; DB 1; Length 779;  
Best Local Similarity 29.6%; Pred. No. 1.5e-53;  
Matches 152; Conservative 105; Mismatches 176; Indels 81; Gaps 16;

QY 135 LDELIDSEPTQVKHMMQVIEPQFORDFISLLPKELALYVLSFLEPKDLQAQOTCRYW- 193  
DB 247 LERLVAMDRSELSDLGTLTKDNKRLDITSLPFEISLKFNYLQFEDIINSLGVSONWN 306  
QY 194 RLAEDNLLWRECKEKEGDEP---LHKRRKVIKPGFIHSPKSAIYRQHRIDTNWR 248  
DB 307 KIIRKSTLWKLKLIENFVSPKGFNSLNKLSQKPKLSQQDRRLSLFLENIFILKNW- 365  
QY 249 RGEKSPK-----VLKGGHDDHVTICLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLUGH 302  
DB 366 ---YNPKEFVPQRTTLRGHMTSVITCLQFEDNVVITGADDKMIRVYDSINKKFLQLSGH 421  
QY 303 TGVWVS-SQMRDNIISGSDRTLKVNAETGECIHTLYGHTSTVRCHLHE-----KRVV 357  
DB 422 DGGWALKVAHGILVSGSDRTLVRVMDIKKGCCTHVFEGHNSVTRCLDVEYKNKIYIV 481  
QY 358 SGRSDATLRVWDI-----ETGQCLHVLGMHVAARCVQYDGR 394  
DB 482 TGSRDNTLHWKLPKESSVDPDHEEHYPLVFTPEENPYFVGLRGMASVTVSGHN 541  
QY 395 RVVSGAYDFMVKVWDPETETCLHTLQHTNRVYSLQFDGTH-----VVSGLDTSIRVWDV 450  
DB 542 IVVSGSYDNTLIVWDVAQMKCLVILSGHTDRIYSTIYD---HERKRCISASMDTTIRIWL 599  
QY 451 E-----TGNC-----IHTLGHOSLSFGMELKNDNLVSGNADSTVKIMDI 490  
DB 600 ENIWNNGECSYATNSASPCAKILGAMVTLQGHATLVGLRLSLDKFLVSAADGSIKWD- 658  
QY 491 KTGQCLQTLQGPKNKHOSAVTCLQFNKFNVTSSDDGTGVKLWDLKTGFEIRNLVLESQGS 550  
DB 659 -ANDYSRKYSHHTNLSAITFFVYSDNLVSGSEN-OFNINLRSKLVHANILKDA--- 713  
QY 551 GGWVWRIRASNTKLVCAGVSRNGTEETKLVLDF 584  
DB 714 -DOIVSVNEFKGTLVAV-EKDG--QSFLEILDF 743

RESULT 5  
US-08-283-917-3  
Sequence 3, Application US/08283917  
Patent No. 5849557  
GENERAL INFORMATION:  
APPLICANT: ADACHI, HIDEKI  
APPLICANT: TSUJIMOTO, MASAFUMI  
APPLICANT: INOUE, KEIZO  
APPLICANT: ARAI, HIROYUKI  
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME  
TITLE OF INVENTION: AND GENE THEREOF  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER &  
ADDRESSEE: NEUSTADT, P.C.

STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/283.917  
FILING DATE: 03-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 209943/1993  
FILING DATE: 03-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5849557man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2292-030-0  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
US-08-283-917-3

Query Match 14.0%; Score 439; DB 2; Length 409;  
Best Local Similarity 30.1%; Pred. No. 7.2e-36;  
Matches 110; Conservative 62; Mismatches 133; Indels 60; Gaps 10;

QY 227 GFHSPPKSAVIRQHRI-----DTNRRGELKSPK-----VLKGHD 262  
DB 48 GLEKKWTSVIRLQKKVMELESKLNKAEKFTSGGPGKQKRPKEWIPRPPEKVALSGHR 107  
QY 263 DHVITCL-QFCGNRIYVSGDDNTLKWSAVTGKCLRTLVTGHTGGVW--SSQMRDNIISG 319  
DB 108 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 167  
QY 320 STDRTLKVNNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377  
DB 168 SADMTIKLWDFQGFECIRTMHGHNDHNVSSVAIMPNGDHIVSASRDKTIKMVEVQTGYCVK 227  
QY 378 VLMGHVAARVCVQ--YDGRVVSGAYDFPMVKVWDPETETCLTTLQGHTRNRYSLQF----- 431  
DB 228 TFTGHRWVRVVRPNQDGTLLIASCNSDQTVRWWVATKECKAELREHEHVVVECSIWAPES 287  
QY 432 -----DGIHVVSGLDTSIRVWDVETGNCIHTLTGHQSILTSGMELKD 473  
DB 288 SYSSISEATGETKSGRPFLLISGSRDKTIKMWDVSTGCLMTLVGHDNVVRGLFHS 347  
QY 474 --NIIYVSGNADSTVKIWDIKTGQCLQTLOGPNKHQSAVTCLQFNKN--FVITSSDDGTVK 529  
DB 348 GKGKFLSCADDKTLRVWDYKKNKCKMKT---NAHEHFTSLDFHKTAPYVVTGVSVDQTVK 404  
QY 530 LWDLK 534  
DB 405 WNECR 409

RESULT 6  
US-08-961-716-3  
Sequence 3, Application US/08961716

Patent No. 5880272  
GENERAL INFORMATION:  
APPLICANT: ADACHI, HIDEKI  
APPLICANT: TSUJIMOTO, MASAFUMI  
APPLICANT: INOUE, KEIZO  
APPLICANT: ARAI, HIROYUKI  
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME  
TITLE OF INVENTION: AND GENE THEREOF  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESS: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961.716  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/283.917  
FILING DATE: 03-AUG-1994  
APPLICATION NUMBER: JP 209943/1993  
FILING DATE: 03-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5880272man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2292-030-0  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
US-08-961-716-3

Query Match 14.0%; Score 439; DB 2; Length 409;  
Best Local Similarity 30.1%; Pred. No. 7.2e-36;  
Matches 110; Conservative 62; Mismatches 133; Indels 60; Gaps 10;  
QY 227 GFHSPPKSAVIRQHRI-----DTNRRGELKSPK-----VLKGHD 262  
DB 48 GLEKKWTSVIRLQKKVMELESKLNKAEKFTSGGPGKQKRPKEWIPRPPEKVALSGHR 107  
QY 263 DHVITCL-QFCGNRIYVSGDDNTLKWSAVTGKCLRTLVTGHTGGVW--SSQMRDNIISG 319  
DB 108 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGHTDSVQDISFDHSGKLLASC 167  
QY 320 STDRTLKVNNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377  
DB 168 SADMTIKLWDFQGFECIRTMHGHNDHNVSSVAIMPNGDHIVSASRDKTIKMVEVQTGYCVK 227  
QY 378 VLMGHVAARVCVQ--YDGRVVSGAYDFPMVKVWDPETETCLTTLQGHTRNRYSLQF----- 431  
DB 228 TFTGHRWVRVVRPNQDGTLLIASCNSDQTVRWWVATKECKAELREHEHVVVECSIWAPES 287  
QY 432 -----DGIHVVSGLDTSIRVWDVETGNCIHTLTGHQSILTSGMELKD 473

Db 288 SYSSISEATGSETKKSGKPGPFLLSSGRDKTIKMWVSTGMCMLTLVGHNDWVRGVLFHS 347  
QY 474 --NILVSGNADSTVKIWDIKTQCLOTLQGNPKHQSAVTCLOFNKN--FVITSDDGTGK 529  
Db 348 GKGFIILSCADDKTLRWVDYKMKCMKTL---NAHEHFTVSLDFHKTAPYVVTGSDQTVK 404  
QY 530 LWDLK 534  
Db 405 WVECR 409  
  
RESULT 7  
US-08-283-917-9  
; Sequence 9, Application US/08283917  
; Patent No. 5849557  
; GENERAL INFORMATION:  
; APPLICANT: ADACHI, HIDEKI  
; APPLICANT: TSUJIMOTO, MASAFUMI  
; APPLICANT: INOUE, KEIZO  
; APPLICANT: ARAI, HIROYUKI  
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
; ADDRESSEE: NEUSTADT, P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/283,917  
; FILING DATE: 03-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 209943/1993  
; FILING DATE: 03-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5849557man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2292-030-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 410 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-283-917-9

Query Match 14.0%; Score 439; DB 2; Length 410;  
Best Local Similarity 30.1%; Pred. No. 7.2e-36;  
Matches 110; Conservative 62; Mismatches 133; Indels 60; Gaps 10;  
  
QY 227 GFTHSPWKSAYINQHRI-----DTNWRGELKSPK-----VLKGHD 262  
Db 49 GLEKKKTSVIRLQKVMESLEKNEAKBEFTSGGPGQKDPKEWIPRPPEKVALSGHR 108  
QY 263 DHVITCL-QFCGNRIYSGDDNTLKVSAVTKCLTLYGHTGGVW--SSQMRDNIILIS 319  
Db 109 SPVTRVIFHPVSVMSVASEDAIKVMDVETGDFERTLKGHTDSVQDISFDSHGKLLASC 168  
QY 320 STDRTLKVWNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRWDIETGQCLH 377

Db 169 SADMTIKLWDFQGFECIRTMHGDHNVSSVAIMPNGDHIVSASRDKTIKMWVGTGYCVK 228  
QY 378 VLGHVAAVRCVQ--YDGRVVSAYDFMWKYVDPEPETCLHTLQHTNRYVSLQF---- 431  
Db 229 TFGHREWRVWRPNODGTLIASCNDQTVRVWVATKECKAELREHEHVVEICISWAPES 288  
QY 432 -----DGIHVVSGLDTSIRVWDVETGNCIHTLTGHSLSGTMELKD 473  
Db 289 SYSSISEATGSETKKSGKPGPFLLSSGRDKTIKMWVSTGMCMLTLVGHNDWVRGVLFHS 348  
QY 474 --NILVSGNADSTVKIWDIKTQCLOTLQGNPKHQSAVTCLOFNKN--FVITSDDGTGK 529  
Db 349 GKGFIILSCADDKTLRWVDYKMKCMKTL---NAHEHFTVSLDFHKTAPYVVTGSDQTVK 405  
QY 530 LWDLK 534  
Db 406 WVECR 410  
  
RESULT 8  
US-08-961-716-9  
; Sequence 9, Application US/08961716  
; Patent No. 5880272  
; GENERAL INFORMATION:  
; APPLICANT: ADACHI, HIDEKI  
; APPLICANT: TSUJIMOTO, MASAFUMI  
; APPLICANT: INOUE, KEIZO  
; APPLICANT: ARAI, HIROYUKI  
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
; ADDRESSEE: NEUSTADT, P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,716  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/283,917  
; FILING DATE: 03-AUG-1994  
; APPLICATION NUMBER: JP 209943/1993  
; FILING DATE: 03-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5880272man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2292-030-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 410 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-961-716-9

Query Match 14.0%; Score 439; DB 2; Length 410;  
Best Local Similarity 30.1%; Pred. No. 7.2e-36;

```

RESULT 10
US-08-899-578-7
; Sequence 7, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/ARK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 212 amino acids
; TYPE: amino acid

```



TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49  
US-08-190-802A-66

Query Match 11.68; Score 364.5; DB 1; Length 514;  
Best Local Similarity 23.08; Pred. No. 3.6e-28;  
Matches 119; Conservative 116; Mismatches 153; Indels 129; Gaps 26;

Qy 155 EPOFORDISLLPRELALYVLSF--LEPKDLL-----QAAQTCRYWRILAED 199  
Db 15 EAQLPRE-VAIIPKDLPNVSIKFOALDTGDNVGGALRVPGAISEKQLEELNQLNGTSDD 73  
Qy 200 NLLWRECKEKG-----IDELHKKRRKVKPKGF-----IHPKMSAYIR 239  
Db 74 PVPYTFESCTIOGKASDPVKITIDITDONLY---SSLIKPGYNSTEDQITLLYTPRAVFKVK 130  
Qy 240 QHRIDTNRRGELKSPVKLVKGDHVVITCLQFC---GNRIVSGDDNTLVKWSAVTGKCL 296  
Db 131 P-----VTRSSSAIAGHGS--TILCSAFAPHTSSRMVMTGAGDNTARIWDCDTQTPM 179  
Qy 297 RTLVGHTGGV---WSSQMRDNIISGSDRTLVKWNNAETGECI-HTLYGHTSVRCM-- 349  
Db 180 HTLVGHYNNVLCVNSPD--GEVIATGSMNDTIIRLWDPKSGCQGLDALRGHSKWITSLW 237  
Qy 350 ---HL-----HEKRVVSGSRDATLRWVDETQGLHVLGMHVAARVRCVYDGRVVY--SGAY 401  
Db 238 EPIHLVKPGSPRLASSKSGDTIKIWDVRSVCQYTMSGHTNSVSCVKGQGGGLLYSGSH 297  
Qy 402 DFVVKVNDPETE-TCLHTPLOGHTNRV-----YSLQ---FDGIH----- 435  
Db 298 DRTVRVNDINSQGRINCILKSHAHVNLSTDYALRIGAFD--HTGKKPSTPEEAQKK 355  
Qy 436 -----VSGSLDTSRVD-VETGNCIHTLGHQSLTSGMELKDN-- 474  
Db 356 ALENYKICKKNGSEENMVTASDDYTFMLNPLKSTKPIARMTGHQKLVNHVAFSPDGR 415  
Qy 475 ILVSGNADSTVKINDIKTGQCLQLOGPNKHOSAVTCLQFNKN--FVITSSDDGTVKLWMD 532  
Db 416 YIVSASFNSIKLMDGRDGRKFISTFRG---HIASVYQVWSSDORLVLVSCSKDTLKVMD 472  
Qy 533 LKTGEFIRNLVTLSESGSGGVVWRIRASNTKLVCAGV 569  
Db 473 VTRKLSVDLPKIKT--KLYVDMSV---DGRKVCSSG 504

RESULT 13  
US-08-188-582-5  
; Sequence 5, Application US/08188582  
; Patent No. 5534410  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.  
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
; NUCLEIC ACIDS ENCODING TAFS AND METHODS OF USE  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,582  
; FILING DATE: 28-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 704 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-188-582-5

Query Match 10.68; Score 332; DB 1; Length 704;  
Best Local Similarity 29.08; Pred. No. 1.2e-24;  
Matches 88; Conservative 61; Mismatches 126; Indels 28; Gaps 13;

Qy 257 VLKGDHDDHVIITCLQFC---GNRIVSGDDNTLVKWSAVTGKCLRTLVGHTGGVWSSQMRDN 314  
Db 373 VLNSHQG--VTCAEISDDSTMLACGFGSSVRKNSLTTPAK-LRTLKD-----ADSLRE- 422  
Qy 315 IISGSDRTLVKWNNAETGECIHTLYGHTSV--RCMHLEKRR-VVSGSRDATLRVMDIET 372  
Db 423 -LDRESADINVRMLDDRSGEVTRSLMGHTGPYRCAPAFAPENMLLLSCSEDSITRLWSLLT 481  
Qy 373 GQCLHVLGMHVAARVRCVQY--DGRRVVSGAYDFVVKVNDPETECLHTLQGHTRNVYSLQ 430  
Db 482 WSCVVYTRGHVVPVMDVRFAPHGYFVSCYDKTARLWATSDNOALRVFVGHLSVDVCVQ 541  
Qy 431 P--DGIHVVSGLDTSIRVWDETGCNHTLGHQSLTSGMELK--DNILYSGNADSTVK 486  
Db 542 FHPNSNYVATGSSDRTVRLWDMNTGQSVRLMTGHHKGSVSSLAFAACRYLASSGVDHNI 601  
Qy 487 IWDIKTGQCLQLOGPNKHOSAVTCLQFNKNFVITSSD--DGTVKLWDLK--TGDFIRNL 542  
Db 602 IWDLSNGSLVITLL---RHTSTVTITFSRGTGLVAAAGLDNNLTLWDFHKVTEYISNH 658  
Qy 543 VTL 545  
; :  
Db 659 ITV 661

RESULT 14  
US-08-646-715-5  
; Sequence 5, Application US/08646715  
; Patent No. 5637686  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:39:35 ; Search time 17.58 Seconds  
(without alignments)  
2552.151 Million cell updates/sec

Title: US-09-328-877A-8  
Perfect score: 3143  
Sequence: 1 MSRPCKPTLHGLVPVDLKS.....SRNGTEETKLLVLDVDMK 589

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1238	39.4	579	2 T22703	hypothetical prote
2	680	21.6	605	2 T38932	probable sulfur me
3	650	20.7	640	2 S49932	MET30 protein - ye
4	641	20.4	703	2 T43557	F-box/WD-repeat pr
5	639.5	20.3	701	2 T16607	hypothetical prote
6	634.5	20.2	518	2 B48088	beta-transducin re
7	620	19.7	775	2 T45136	WD repeat protein
8	618.5	19.7	779	2 S56245	cell division cont
9	582.5	18.5	1356	2 T18521	beta transducin-li
10	576.5	18.3	650	2 T46660	sulfur controller-
11	533	17.0	506	2 T50211	WD-repeat protein
12	446.5	14.2	1146	2 A55532	myosin-heavy-chain
13	439	14.0	410	2 S48052	platelet-activatin
14	436.5	13.9	409	2 S36113	Lis-1 protein - hu
15	410.5	13.1	317	2 T46032	WD-40 repeat regul
16	403	12.8	267	2 S62507	hypothetical trp-a
17	396	12.6	376	2 T19265	hypothetical prote
18	392.5	12.5	333	2 G85034	probable WD-repeat
19	378.5	12.0	502	2 T41148	trp-asn repeat con
20	376.5	12.0	422	2 A56640	CDC4 repeat unit-c
21	376	12.0	501	2 T27513	hypothetical prote
22	365	11.6	515	2 S19487	hypothetical prote
23	361	11.5	395	2 T23117	hypothetical prote
24	360	11.5	1693	2 S76086	beta transducin-li
25	353.5	11.2	714	2 S56893	hypothetical prote
26	336	10.7	777	2 T41075	hypothetical WD-re
27	333.5	10.6	494	2 T19550	hypothetical prote
28	330	10.5	486	2 S49820	PRL1 protein - Ara
29	326	10.4	704	2 S33263	transcription init

WD repeat protein  
probable pleiotrop  
WD-40 repeat regul  
hypothetical prote  
WD repeat protein  
protein TipD - sll  
PRL2 protein - Ara  
neuronal different  
hypothetical prote  
zinc finger protei  
probable GTP-bindi  
hypothetical prote  
WD40-repeat protei  
WD-40 repeat regul  
TATA box-binding p  
beta-transducin -

ALIGNMENTS

RESULT 1  
T22703  
hypothetical protein F55B12.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T22703  
R:Simms, M.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z19602  
A:Accession: T22703  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-579 <NII>  
A:Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GN00023; CESP:F55B12.3  
A:Experimental source: clone F55B12  
C:Genetics:  
A:Gene: CESP:F55B12.3  
A:Map position: 5  
A:Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 39.4%; Score 1238; DB 2; Length 579;  
Best Local Similarity 45.6%; Pred No. 9.8e-90;  
Matches 262; Conservative 85; Mismatches 162; Indels 66; Gaps 11;  
QY 55 DHGSEVRSFSLGKKPC-----KVSEYTSFTTGLVPCSATPTTFGLRAANGGQQRRTSV 110  
Db 23 DNGEE-SSYSNGSSSSYNADKLSSSRPLQHKLDLSAPSRNNDLNPR----- 68  
QY 111 QPPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVKHMQVIEPQFORDFISLLPKEL 170  
Db 69 -----VEHLIALFKDLSAEQMDAFTRLQESNMTNIQRLRAITIEPHFQDFLSCLPV 123  
QY 171 ALVLSFLEPKDLQAQTCRYWRILAEADNLNLRKCKEE----- 210  
Db 124 GKKILHNLGYDLLKVAQVSKNWKLLISEIDKWKSLGVVEEFKHPDPDPTDRVTGAWGCTAI 183  
QY 211 --GIDEPLHI-----KRRKVIKPG--FIHSPMKSAIYIRQHRIDTNWRRGELKSPKVL 258  
Db 184 AAGVTIPDHIQPCDLNVHRFLKQKFGDIFERAADKSRVLRADKIEKNWNNANFIMSASV 243  
QY 259 KGHDDHVTICLQFCGNRIYVSGDDNTLKVSAVTKCLRTLVGHTGVVSSQNRD--NII 316  
Db 244 RGHEDHVTICMIHDDVLVTGSDNTLKYCIDKGEVMTLVGHTGVVWTSQISQCGRYI 303  
QY 317 ISGSTDRTLKVNAETGECIHTLYGHTSVRCMHLKEKRVWSGSRDRLRVWDIETGQCL 376  
Db 304 VSGSDTRTVKVNSTVDGSLHTLQGHSTVRCMAMAGSILVTGSGRDLTKRVWVDSGRHL 363  
QY 377 HVLGMHVAARVCQVQDGRRVVSGADFMVKVMDPETETCLHTLQGHTRNVYSIQFDGIIH- 435  
Db 364 ATLHGHHAAVRCVQDGTIVVSGGYDFTVKIWNNAHTGRCIRTLTGHNRVYSLLFESERS 423

Qy 436 -VVSGSLDTSIRVWDV---ETGNCIHTLTGHQSILTSCMELKNLIVSGNADSTVKTWIDIK 491  
Db 424 IVCSGSGLDTSIRVWDFTRPGEQCVALLOHGTSLTSGMLGRNLIYSCNADSHRVYWDIH 483  
Qy 492 TGOCQLTLQGPNKHQSVAVTCLQ-FNKNFVITSSDDGTGVKLKWLDTKEFTLRNLVLTESSGS 550  
Db 484 EGTCVHMLSG--HRSAITSLQFWRNVATSDDDGTGVKLKWLDERGALIRDLVLTDGGN 540  
Qy 551 GGVVRIIRASNTKLVCAGVRNGTEETKLLVLDPD 585  
Db 541 GCCIWRLCSTSTMACAVGSRNTEETKVILLDFD 575

RESULT 2  
T38932  
probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 26-May-2000  
C:Accession: T38932  
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL data Library, April 1997  
A:Reference number: Z21818  
A:Accession: T38932  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-605 <BAD>  
A:Cross-references: EMBL:Z94864; PIDN:CAB08168.1; GSPDB:GN00066; SPDB:SPAC57A10.05c  
A:Experimental source: strain 972h-; cosmid c57A10  
C:Genetics:  
A:Gene: SPDB:SPAC57A10.05c  
A:Map position: 1  
C:Superfamily: unaligned WD repeat proteins; WD repeat homology

Query Match 21.6%; Score 680; DB 2; Length 605;  
Best Local Similarity 29.8%; Pred. No. 1.1e-45;  
Matches 164; Conservative 93; Mismatches 181; Indels 112; Gaps 17;

Qy 51 KRKLHDGSEYRSFLGKKPKCKVSEYTSTTGVLPCSATPTTFGLRAANGQGORRITSV 110  
Db 31 KKEGDQISVSFAFNIS-----SMHNLSGL-----SEKSQRQEAV 66  
Qy 111 QPPYGLEWLMKFQSWGPGPKLLALDELIDSCPTQVKHMVIEPQFORDFISLLPKEL 170  
Db 67 -----W-AAPEASEASCERKLAQILNNCSSLSLSPASTLDSLRLDLSLLPVEI 117  
Qy 171 ALVLSLEPKDLLQAOTCRYWRIAEONLLWREKCE----- 209  
Db 118 SFRILSFDARSLCOAAQSVKHWKELADDVIWHRMCEQHINRKCEKCGWGPLLERNTL 177  
Qy 210 -----EGID--EPLHKRRKY-----IKP----- 226  
Db 178 YAAKASTOKRYRLTKRGVDQAHESSPVKAKLDPTYTSSNEETISSVAPPSPNSDSKFF 237  
Qy 227 -GFTHSPWKSAIYIRHRIDTNWRRELKSPVKLGHHDDHVITCLQFCGNRIYVGSDDNLT 285  
Db 238 LPEKTREPWEKVEYAERCVRVECNWRHGCRQ-VLSGHSDCGM-CLQLVRNILASGYDATI 295  
Qy 286 KVNSAVTGKCLRILTGHGTGWSSQMRDNIISGSTDRTLKYVNAETGCIHTLYGHTST 345  
Db 296 RLWNLATFQQVALLEGHSSGVTCLQPDQCKLISGSMDKTIRIWNVRTSECISILHGTDS 355  
Qy 346 VRCHHLHEKRVSISRDATLRVWDIETGCQLHLVLMGHVAAVRCVQY--DGRVYVSGAYDF 403  
Db 356 VLCLTDFDSTLLVSGSADCTVKLMHFSGGKRI-TLRGHTGPVNVSRIIRDGLVLVSDDS 414  
Qy 404 MVKWVPDETCLTHTLQGHNTNRVYSIQFDGIHWSGSLDTSIRVWDVETGNCIHTLTGHQ 463  
Db 415 TIKIWSLETNTCLHTTSAHIGPVOSIALADSRFLFCSLDGTIKQNDIEKKKCIVHTLFCH- 473  
Qy 464 SLTSGMEL-KDNI-LVSGNADSTVKTWIDTKTGCCLOTLOGPNKHQSVAVTQCFQNKFNIT 521

Db 437 FSCSDTTIRMDIRNSCLKVFRGHVGQVQKIIPIITIKIDVENLATDNTSDGSSPQDDPT 496  
 QY 430 QFDG-----IHVSGSLDTSIRWVDVETGNCIHTLTGHOSLTS 467  
 Db 497 MTGDADESOTPSNEQETVLDENTIPYTHILLSQGLDNTIKLMDVKTKCINTQFGRHVEGVW 556  
 QY 468 GMELKNILVSGNADSTVKIWDIKTQCLOQT-----LOGPNKHQ-----SAYTCLQ 513  
 Db 557 DIAADNFRITSGSHDGSIKVMDLQSKCMHTFNGRRRLQRETQHTQTQSLGDKVAPIACVC 616  
 QY 514 FNKNFVITSSDDGTVKLW 531  
 Db 617 IGDSECFSGDEFCVKMY 634

RESULT 4  
 T43557  
 F-box/WD-repeat protein pop2 - fission yeast (Schizosaccharomyces pombe)  
 N:Alternate names: proteolysis factor sudlp  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T43557; T38794; T43798  
 R:Wolf, D.A.; Jackson, P.K.  
 submitted to the EMBL Data Library, December 1997  
 A:Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th  
 A:Reference number: Z22576  
 A:Accession: T43557  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-703 <WOL>  
 A:Cross-references: EMBL:AF038867; PIDN:AAB95480.1  
 A:Experimental source: strain h- 972  
 R:Gentles, S.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z21812  
 A:Accession: T38794  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-703 <GEN>  
 A:Cross-references: EMBL:Z98602; PIDN:CAB11275.1; GSPDB:GN000066; SPDB:SPAC4D7.03  
 A:Experimental source: strain 972h-; cosmid c4D7  
 R:Jallepalli, P.V.; Tien, D.; Kelly, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998  
 A:Title: Sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1 proteins fo  
 A:Reference number: Z22686; MUID:98318628  
 A:Accession: T43798  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-703 <JAL>  
 A:Cross-references: EMBL:AF064515; NID:g32933382; PIDN:AAC39496.1; PID:g32933383  
 A:Gene: pop2; sud1; SPAC4D7.03  
 A:Map position: 1  
 C:Function:  
 A:Description: required to prevent spontaneous re-replication

Query Match 20.4%; Score 641; DB 2; Length 703;  
 Best Local Similarity 27.3%; Pred. No. 1.7e-42;  
 Matches 178; Conservative 111; Mismatches 248; Indels 116; Gaps 17;  
 QY 9 LNHGLVPVLDKSAKEPL-----PHQTMKIFSIITIAQ-----GLP-----F 45  
 Db 85 LKHSCTPSELRKFNESIENVYKCLDHPDPSVGDGFSILVFPQRFNLYSHSLPPKIIS 144  
 QY 46 CRRMRKRLDHGSEVRFSGLGKPKCVSEYTTGLVPCSATPTFGDLRAANGQOQRR 105  
 Db 145 IDRNRIKLDNISNSDNFPPSPKVDNTNTVSPGSKPISE-----DLEDNLNLSIVO- 197  
 QY 106 RITSVOPPTGLQEWLKMFMOSGPEKLLALDELIDSCPTQVKHMMQVIEPQFQDFISL 165  
 Db 198 --TFEDLPBGIOY-----AFFQLLRSCNQSRMLLNECEPLLKRDILSN 241

QY 166 LPKELALYVLSFLEPKDLLOAAOATCRYW-RILAEADNLLWRECKCKEGI---DEPLHIKRR 221  
 Db 242 LPSISVQSILLNLDIHSFLSCLVSTWNRILDVHTSYWKHMFSLGFOQINENDWKYANP 301  
 QY 222 KVIKPGFIHSPWKS-----AYIRQHRIDTNWRRGELKSPKVLKGH-----DDHVITCL 269  
 Db 302 NLNRPPFLHNDQISDDYFPEIFKRHFLN-----RKRWLFPSPISPPSHLSFPIHVPNEMITSL 357  
 QY 270 QFCGNRVSGSDNTLKVMSAVTGKCLRTLVGHTGGWSSQMRDNIISGSTORTLKVWN 329  
 Db 358 LHKDRIITTSGSGTQIHNATGVLEARLEGHKEGVAVKIHENTILVSGSIDKTVRVWN 417  
 QY 330 AETGECIHTLYGHTSVRCMHL-----HEKRVVSGSRDATLRWDI--- 370  
 Db 418 IEKAKCTHIFRGHISIIRCLLEILVPSRLRHGVEIVPEPOPIYVSGSRDHTLRWKL PKN 477  
 QY 371 -----ETGQCLHLVLMGHAARVRCVQYDGRVRYVSGAYDPMKVPWPETE 413  
 Db 478 TDPYLPDNTNSIDRWEKNPYFVHTTLIGHTSVRTISGYDILVSGYSSIRIRWVSTG 537  
 QY 414 TCLHTLQGHTRNVYSLOFDGIH--VYSGSLDTSIRWVDVETGNCIHTLTGHQSLTSGMEL 471  
 Db 538 ECLYHLRGHSLRIYSVLYEPERNICISGSMKDSIRVMDLSTGTCKYVLEGHDAFVTLNV 597  
 QY 472 KDNILVSGNADSTVKIWDIKTQCLOQTLOGPNKHQSAVTCLOQKNFVITSSDDGTVKLW 531  
 Db 598 FQNRLLSGADSTIRIWDLNTGKPLMVLPSNSGYISFV-----SDEHKIISGNDGSKLW 653  
 QY 532 DLKTGEFIRNLVLTESGGGVVVRIRASNTKLVCVAGSRNNGTEETKLVLDF 584  
 Db 654 DVRTGKLLRFLT-----DLTKIWHVDFEDAMRCVAAV---QRDDQAYLEVINF 698

RESULT 5  
 T16607  
 hypothetical protein K10B2.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16607  
 R:Miller, N.  
 submitted to the EMBL Data Library, June 1995  
 A:Description: The sequence of C. elegans cosmid K10B2.  
 A:Reference number: Z18545  
 A:Accession: T16607  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-701 <MIL>  
 A:Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAA68258.1; CESP:K10B  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:K10B2.1  
 A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 20.3%; Score 639.5; DB 2; Length 701;  
 Best Local Similarity 33.5%; Pred. No. 2.2e-42;  
 Matches 154; Conservative 80; Mismatches 167; Indels 59; Gaps 12;  
 QY 126 WSGPEKLLALDELIDSCPTQVKHMMQVIEPQFODFISLLPKELALYVLSFLEPKDLQ 185  
 Db 83 WSEHEQLDFMDKIVHRLSHVQLGKVDNFRPMLORDFISNLPALHVELIL-FNVNSDSLK 141  
 QY 186 AAQ-----TC-----RYWRIILAEADNL-----LW-----REKCKE 209  
 Db 142 SCEEVSTSNRCALARGQHWKLLTEKNVRSDSLWGLSEKQWMDKFLNISRMSVRRICEK 201  
 QY 210 EGIDEPHLHKRRKVKPGFIHSPWKSAY---TRQ-HRIDTNWRRGELKSPKV-LKGDH 264  
 Db 202 FNYD--VNIRKDKLDQILMHVYFKLYPKIIRIHDNIDNNKRGNYKMTIRNQCSENSK 259  
 QY 265 VITCLOFCGNRVSGSDNTLKVMSAVTCCKLRTLVGHTGGWSSQMRDNIISGSTORT 324  
 Db 260 GVICLOYDDKIVSGLRDNTIKIWRKDYCSRIILSGHTGSLVLCIQYDNRVILSGSSDAT 319



366 ---YNPKFPQRTTLRGHMTSVITCLQFDNVDYVITGADDKMIRVYDSINKKFLQLSGH 421  
 QY TGGVWS-SQMRDNIILSGSDTRILKYVNAETGECIHTLYGHTSTVRCMHLE-----KRVV 357  
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 422  
 422 DGGVWALKYAGGILVSGSDTRVRVMDIKKGCTHFVKGHNSTVRCCLDIVEYKNIKYIV 481  
 QY SGRSDATLRVMDI-----ETGQCLHVLHMGHVAARCVQYDGR 394  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 482  
 482 TGSRDNTLHWKLPKESSVPDHGEEHDYPLVFHTPEENPYFGVGLRGHMASVTVSGHN 541  
 QY RVSGAYDFWVKVWDPEETETCLHTLQGHTRRVYSLQFDGIH-----VYSGSLDTSIRVDV 450  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 542  
 542 IVSGSDNTLLIYVMDVAQMKCILYSGHTDRIYSTIYD--HERKRCISASMDTTIRIWL 599  
 QY E---TGNC-----IHTLGHOSLTSGMELKDNILVSGNAUSTVKIWDI 490  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 600  
 600 ENIWNNGECVYATNSASPCAAILGAMVTLQGHATLVGLLRSLDKFLVSAADGSRGWD- 658  
 QY KTGOCLQTLOQPNKHOSAVTCLQFNKNFVITSSDDGTVKLMDLKTGFEIRNLVLTLES 550  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 659  
 659 -ANDYSKFSYHHTNLSAITTFVSDNILVSGSEN-QFNIYNLRSKGLVHAMILKDA--- 713  
 QY 551 GGVVWRIRASNTKLVCVAGSRNGTEETKLVLDF 584  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 714  
 714 -DQIWSVNFKGKTLVAAV-EKDG-QSFLFILD 743

RESULT 9  
 Tl8521  
 beta transducin-like protein - Podospora anserina  
 C:Species: Podospora anserina  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: Tl8521  
 R:Saupé, S.; Turcq, B.; Begueret, J.  
 A:Title: A gene responsible for vegetative incompatibility in the fungus *Podospora anserina*  
 A:Reference number: Z18944; MUID:96009891  
 A:Accession: Tl8521  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1356 <SAU>  
 C:Genetic references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1  
 A:Gene: het-el  
 A:Introns: 761/3

Query Match 18.5%; Score 582.5; DB 2; Length 1356;  
 Best Local Similarity 28.9%; Pred. No. 1.8e-37;  
 Matches 169; Conservative 85; Mismatches 193; Indels 137; Gaps

QY 113 PTGLQEWLKM-----FQSWSGPEKLL-----ALDELIDSCEPTQVKH 149  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 590  
 590 PESLEEIVKLCGSFLIIRERTVYFVHQSADKFLGTTASDKASKNASQAEFELVFTGID 649  
 QY 150 M-----MOVIEPQQRDFISLLPKELALYVLSLEPKDLLQAQTCRYW----- 193  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 650  
 650 VSYIIFWRSNLVMSQKLRRDIYCLNAPGFLIDNVRPDPPLATVRYSCIYWIHDRLDIV 709  
 QY 194 -----RIEADNLLWR-----KCKEEDID-----BPL--HKRRK- 222  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 710  
 710 SSTSSKVVHLLDDGDTHREFTKTYLWLEALSLLRALPEGINATROLESGLGHTIRGL 769  
 QY 223 --VIKPGF-----IHPWKSAYIRQHR-----IDTN 246  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 770  
 770 IAIVRDGYRFALSYRMIEKAPLQAYTSALVAPTDSMIKKIFKKEPGWISTISVWEA 829  
 QY 247 WRRGELKSPKVLGHDDHVT-CLQFCGNRIYVSGSDNTLKYWSAVTGKCLTLVGHGTG 305  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 830  
 830 WN-----ACTGTLEGHGSSVLSVAFSADGORVASGSDDKTIKIWDIASGTGTGLEHGG 885  
 QY 306 VWS---SQMRDNIILSGSDTRTLKVVNAETGECIHTLYGHTSTVRCMHLE---KRVVSGS 360  
 Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Db 886 VWSVAFSPDRER - VASGSDDDKTIKIWDASGCTCTQTLEHGGRVQSVAFSPDQORVASGS 944  
Qy 361 RDAIRVMDIETGQCLHVLGMHGAARVCQY - DGRRVVSGAYDFMKVMDPETETCLHT 418  
Db 945 DDHIKIWDASGCTCTQLEHGSSVLSVAFSPDQORVASGSKTIKIWDASGCTCTQT 1004  
Qy 419 LOGHTNRVYSLOF - DGIHVVSGLSDTSIRVMDVETGNCIHTLTGH - - - - - QSLTSGMELK 472  
Db 1005 LEGHGSVWSVAFSPDQORVASGSDDKTIKIWDASGCTCTQLEHGSGVQSVVFSFD - 1062  
Qy 473 DNILVSGNADSTVKIWDIKTQCQLOTLOGPNKHOSAVTCLQFNK - - - - - FVITSSDDGTVKL 530  
Db 1063 GORVASGSDHDTIKIWDASGCTCTQLEG - - - - - HGDSVWSVAFSPDQORVASGSDGTIKI 1119  
Qy 531 WDLKTGFIRNLVLTESGGSGVWVRIRASNTKLCVACVSRNGT 574  
Db 1120 WDAASGCTCTQ - - - - - TLE - - - - - GHGWSVAFSPDQORVASGSDGT 1158

RESULT 10  
T46660  
sulfur controller-2 protein [Imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
C:Accession: T46660  
R:Kumar, A.; Paietta, J.V.  
Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995  
A:Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encodes a  
A:Reference number: T23121; MUID:95241499  
A:Accession: T46660  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-650 <KUM>  
A:Cross-references: EMBL:U17251; NID:g806757; PIDN:AAA68968.1; PID:g806758  
C:Genetics:  
A:Gene: scon-2  
A:Map position: 3  
A:Introns: 75/3; 319/1; 354/1  
C:Function:  
A:Description: negatively regulates sulfur structural gene expression  
A:Note: scon-2+ expression is dependent on cys3 function and the binding of cys3 to the  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 18.3%; Score 576.5; DB 2; Length 650;  
Best Local Similarity 26.3%; Pred. No. 1.9e-37;  
Matches 155; Conservative 84; Mismatches 177; Indels 173; Gaps 17;  
Qy 106 RITSVQPTGLQEWLKFQSWGPEKLLALDELIDS - CEPTQVXHMVQVIEPQFQRFIS 164  
Db 71 KLTSADQOAVTHVW - SLFSAAPARHRDLMLQILSQLCFP - QLSFVSREVREALKIDFIS 128  
Qy 165 LLPKALALVLSFLBPKDLLAAOQTCRYWRILAEIDLW - - - - - REKCKEIGIDPL 216  
Db 129 ALPVELAQVLCYLDVSLTKAAQSQWRITLADSDAVVWRCQHVNRKCTKCGWGLPL 188  
Qy 217 - - - - - HIKRRKVIKPG - - - - - 227  
Db 189 LERKLRNTRQRLQAKGPGQVRTELADSHDSQDRSVNQHGKRPAAEAEEEDPIKRRQC 248  
Qy 228 - - - - - FIHSP - - - - - WKSAYIRQHRIDTNRRGELKSPKVLKGDHVDVITCLOFCGNR 275  
Db 249 MAAEASRAVTPKTRSNKAVYRDRQVSYNKNRSRYKL - SVLKGHENG - TCLQLDDNI 306  
Qy 276 IVSGSDDTNLKWSAVTCKCLRTLVGHGGVWSSQMRNIIISGSDTDLKVNAAETGEC 335  
Db 307 LATGSDTITKIWNITETECIRTLVGHAGRALQFDKSLISGSLDITIKWNHHTGEC 366  
Qy 336 IHTLVGHSTVRCMHLEKRVVSGSDATLRVMDTETGQCLHVLGMHGAARVCQYD - - G 393  
Db 367 LSTFAAHTDSVLSVHFDGHLASGSDKTKIFDPNSKE - TYCLKGHSWNSHTVDLKS 425  
Qy 394 RRVVSGAYDFMKVMDPETETCLHTLOGHTNRVYSLOF - - - - - 431

Db 426 RTVFSASDDTTIKLWDLDTQVIRTYEGHVHQVQLLTPPEYEPDEEVLNGASQDNDA 485  
Qy 432 - - - - - DGIIH - - - - - 435  
Db 486 MSVSSGGSGSSMSHAQIERAGSPGSHSSHNLLPSSLPSSGDEEDVRHLYGSAFVADESRP 545  
Qy 436 - - - - - VVSGSLDTSIRVMDVETGNCIHTLTGH - - - - - QSLTSGMELKONI - LVSGNADSTV 485  
Db 546 LPFRYFMGTGGLDSFMRLWDSATGRCLRTLFGHLEGVNSLAG - - - - - DTIRVISGANDGMV 600  
Qy 486 KIMDIKTGQCLOTLOGPNKHOSAVTCLQFNKFNVTSSDDGTVKLWDLK 534  
Db 601 KTWEPRSKCDATVTG - - - - - HCGPVTGCLSDSLMASGSEDGTIRLHSEK 646

RESULT 11  
T50211  
WD-repeat protein [Imported] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
C:Accession: T50211  
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.;  
submitted to the EMBL Data Library, January 2000  
A:Reference number: Z25046  
A:Accession: T50211  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-506 <MCD>  
A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05  
A:Experimental source: strain 972h(-); cosmid c30  
C:Genetics:  
A:Gene: SPAC29E6.01; SPDB:SPAC30.05  
A:Map position: 1  
A:Introns: 43/1; 74/3  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 17.0%; Score 533; DB 2; Length 506;  
Best Local Similarity 25.8%; Pred. No. 3.5e-34;  
Matches 140; Conservative 91; Mismatches 204; Indels 108; Gaps 14;  
Qy 55 DHGSEVRSFSLGKKPKVSEYT - - - - - STTGLVPCSATPTTFGDLRAANGQGOQRRRITS 109  
Db 5 DNGKNVVS - - - - - KVSDLTSCDSFSTSPVPC - - - - - LNPLSHENNRIDL 44  
Qy 110 VQPTGLQEWLKFQSWGPEKLLALDELIDSCEPTQVXHMVQVIEPQFQRFISLLPKE 169  
Db 45 IR - - - - - DLLASLSKEGVVAVYNHVRSLLTFTDTEVFPEE 79  
Qy 170 LALVLSLEPKDLLQAAOQTCRYWRILAEIDLW - - - - - IDEPL - - - - - 216  
Db 80 VSLRVFSYLDQDLCKCKLMSKRWKRLLEDPCINKALYMQKGFVNVNLFNEFEAWRRTH 139  
Qy 217 - - - - - HIKRRKVIKPG - - - - - WKSAYIRQHRIDTNRRGE - 251  
Db 140 KFPQPRFENFLKQNNIGPYGTMLPQOIFDNGRPLNLSYLYKEHAHLDSNNRRHRF 199  
Qy 252 - - - - - LKSPKVLKGDDHV - - - - - ITCLQFCGNRIVSGSDNTLKVMSAVTGKCLRTLVG 301  
Db 200 LVSTFNFNISIRFPADQDFRATLDSVYCYQYDDEINWVSGSKORTSVWVDVNSRFILYKLG 259  
Qy 302 HTGGVWSSQM - - - - - RDNIIISGSDTDLKVNAAETGECIHTLVGHSTVRCMHLEKRVVSG 359  
Db 260 HSGVLCCLDFCRRRLVYSGSDSTIIIDWQNRRLPKVYEGHTDNLVGVVSVSENIIS 319  
Qy 360 SRDATLRVMDIETG - - - - - CLHVLGMHGAARVCQYDGR - - - - - VVSGAYDFMKVMDPETE 413  
Db 320 SRDHTARVRLDATSPAECAMHVLGRHLASVNSVQSKTGLIVTASSDRLTWTDTTG 379  
Qy 414 TCLHTLOGHTNRVYSLOFDTGTHVVSGLDTSIRVMDVETGNCIHTLTGHQSLTSGMELK 473  
Db 380 HCIRIIHAHQGIACAQYNGKFIIVSGSSDLTIRIFEASSGKLLRMLOQHEDLIRTVRND 439



S36113  
LIS-1 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 26-May-2000  
C:Accession: S36113  
R:Reiner, O.; Carozzo, R.; Shen, Y.; Wehnert, M.; Faustinella, F.; Dobyns, W.B.; Cas  
Nature 364, 717-721, 1993  
A:Title: Isolation of a Miller-Dieker lissencephaly gene containing G protein beta-su  
A:Reference number: S36113; MIMID:93361119  
A:Accession: S36113  
A>Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-409 <RAS>  
C:Genetics: LIS-1  
A:Gene: LIS-1  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
F:103-136/Domain: WD repeat homology <WD1>  
F:145-178/Domain: WD repeat homology <WD2>  
F:187-220/Domain: WD repeat homology <WD3>  
F:229-262/Domain: WD repeat homology <WD4>  
F:333-366/Domain: WD repeat homology <WD5>  
F:373-408/Domain: WD repeat homology <WD6>

Query Match 13.9%; Score 436.5; DB 2; Length 409;  
Best Local Similarity 34.0%; Pred. No. 1.1e-26;  
Matches 99; Conservative 54; Mismatches 103; Indels 35; Gaps 7;  
  
QY 276 IVSGSDNTLKWASAVTGKCLRTLVGHGGVW--SSOMRDNIISGTSRDLTKVWNAETG 333  
DB 122 MVSASEDATIKWVDYETGDFERTLKGHTDSVQDISFDHSGKLLASCADMTIKLWDFGFG 181  
  
QY 334 ECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLHLVLMGHVAAVRCVQ- 390  
DB 182 ECIRTMHGHDRHNVSSVAIMPNGDHIVSASRDKTIKMWEEVQTGYCVKFTTGHREWVRVVRP 241  
  
QY 391 -YDGRVVSGAYDFMVKVWDPETETCLHTLQHTNRYSLQF----- 431  
DB 242 NODGTLIASCSNDQTVRVVWVATKECAELREHEHVVEICSWAPESSYSSISEATGSETK 301  
  
QY 432 ----DGIHVVSGLDTSIRVWMDVETGNCIHTLTGHQSLTSGMELKD--NILVSGNADSTV 485  
DB 302 KSGKPGFLLSGSRDKTIKMDVSTGMLTLVGHDNVVRGVLFHSGGKFKILSCADDKTL 361  
  
QY 486 KTWIDIKTGQCLQTLQGPKNKHOSAVTCLQFNKN--FVITSSDDGTVKLWDLK 534  
DB 362 RVWDYKNKRCMKTL---NAHEHFTVSLDFHKTAPVVTGSDVQTVKVVWECR 409  
  
RESULT 15  
T46032  
WD-40 repeat regulatory protein tup1 homolog - Arabidopsis thaliana  
N:Alternate names: protein T16K5.10  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 26-May-2000  
C:Accession: T46032  
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23015  
A:Accession: T46032  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-317 <RIE>  
A:Cross-references: EMBL:AL132965  
A:Experimental source: cultivar Columbia; BAC clone T16K5  
C:Genetics:  
A:Map position: 3  
A:Introns: 222/3  
A:Note: T16K5.10  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 13.1%; Score 410.5; DB 2; Length 317;  
Best Local Similarity 31.9%; Pred. No. 8.4e-25;  
Matches 98; Conservative 66; Mismatches 106; Indels 37; Gaps 12;  
  
QY 256 KVLKGHDDHVITCLOFC--GNRIVSGSDNTLKWWSAVT-----GKCLRTLVTGHGGV-- 306  
DB 18 QLTLSH-NRAVSSVFSSDGRLLASASADKTIYTIINTINDPIAEPVQFEETGHENGISD 76  
  
QY 307 --WSSOMRDNIISGTSRDLTKVWNAETGECIHTLYGHTSTVRCMHLH--EKRVVSGSRD 362  
DB 77 VAFSSDAR--FIVSASDDTKLWMDVETGSLIKTLIGHTNFAFCVFNPNQSNMIVSGSFD 134

QY 363 ATRLRVWDIETGQCLHLVLMGHVAAVRCVQY--DGRRVVSGAYDFMVKVWDPETETCLHTLQ 420  
DB 135 ETVRIWDVTTGCKLVLPAAHSDPVTAVDFNRDGSLLIVSSSDGLCRINWDSGTGHCVKTLI 194  
  
QY 421 GHTNRVYS-LQF--DGIHVVSGLDTSIRVWMDVETGNCIHTLTGHQ-----SLTS 467  
DB 195 DDENPPVSFVRFPSPNGKFLVGTLDNTLRLNWNISSAKFLKTYTGHVNAQYCISSAFSVTN 254  
  
QY 468 GMELKDNILVSGNADSTVKIWDIKTGQCLQTLQGPKNKHOSAVTCLQFNKNFVITSSDDGT 527  
DB 255 GKR-----IVSGSEDNCVHMMELNSKLLQKLEGTETVMNVAC-HPTENLIASGLDKT 308  
  
QY 528 VKLMDLK 534  
DB 309 VRIWTPQK 315

Search completed: September 25, 2001, 14:42:04  
Job time: 149 sec



**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1238	39.4		579	1	SE10_CAEEL	Q93794 caenorhabdi
2	725.5	23.1		684	1	CC4_CANAL	P53699 candida alb
3	680	21.6		605	1	POF1_SCHPO	P87053 schizosacch
4	650	20.7		640	1	MT30_YEAST	P39014 saccharomyc
5	645	20.5		678	1	SCOB_EMENI	Q00659 emericeella
6	641.5	20.4		542	1	FW1B_HUMAN	Q9ukb1 homo sapien
7	641	20.4		703	1	POP2_SCHPO	014170 schizosacch
8	639.5	20.3		701	1	YSS1_CAEEL	Q09990 caenorhabdi
9	638	20.3		605	1	FW1A_HUMAN	Q9Y297 homo sapien
10	634.5	20.2		518	1	TRCB_XENLA	Q91854 xenopus lae
11	620	19.7		775	1	POF1_SCHPO	P87060 schizosacch
12	618.5	19.7		779	1	CC4_YEAST	P07834 saccharomyc
13	582.5	18.5		1356	1	HET1_PODAN	Q00808 podospora a
14	576.5	18.3		650	1	SC02_NEUCR	Q00127 neurospora
15	516	16.4		732	1	KWBH_DICDI	P90648 dictyosteli
16	461	14.7		362	1	WRB5_HUMAN	Q9ugp9 homo sapien
17	460	14.6		361	1	WDS_DROME	Q993j8 drosophila
18	446.5	14.2		1146	1	KMHA_DICDI	P42527 dictyosteli
19	439	14.0		409	1	L1S1_HUMAN	P43034 homo sapien
20	439	14.0		409	1	L1S1_MOUSE	P43035 mus musculus
21	438	13.9		409	1	L1S1_BOVIN	P43033 bos taurus
22	403	12.8		267	1	YAF1_SCHPO	Q09855 schizosacch
23	396	12.6		376	1	YK4_CAEEL	Q17963 caenorhabdi
24	392.5	12.5		422	1	FBW2_HUMAN	Q9ukt8 homo sapien
25	383	12.2		742	1	PKWA_THECU	P49695 thermomonos
26	376.5	12.0		422	1	FBW2_MOUSE	Q60584 mus musculus
27	376	12.0		501	1	YH92_CAEEL	Q23256 caenorhabdi
28	365	11.6		515	1	YCW2_YEAST	P25382 saccharomyc
29	361	11.5		395	1	Y2LL_CAEEL	Q93847 caenorhabdi
30	360	11.5		1693	1	Y163_STNY3	Q55563 synechocyst
31	355	11.3		444	1	NUDF_EMENI	Q00664 emericeella
32	353.5	11.2		714	1	YJ12_YEAST	P47025 saccharomyc
33	342.5	10.9		473	1	PR11_SCHPO	013615 schizosacch

```
Query Match          39.4%; Score 1238; DB 1; Length 579;
Best Local Similarity 45.6%; Pred. No. 9.4e-90;
Matches 262; Conservative 85; Mismatches 162; Indels 66; Gaps 11;

QY 55 DHGSEVRSFSLGKFCPC----KYSEYTSITGLVPCSATPTTFGLRAANGQGOORRITSV 110
Db 23 DNGEE--SSYSGSSSYNADKUSSRPQHKLDLSAPSRNNDLNP-----68

QY 111 OPTGLQEWLKMFSQWSPEKLLALDELIDSCPTQVQVHMVIEPQFORDFISLLPKEL 170
Db 69 -----VEHLIALFKDLSAEOQDAFRLLOESNMNINQLRAIIEPHQFORFSLCLPVEL 123

QY 171 ALYVLSFLEPKDLQAQTCRVWRILAEADNLLWRKCKEE-----210
Db 124 GKMILHNLGTGYDLLKVAQVSKNWKLISETDKLWKSGLVGEFHKHDPDTRVTGAQGTAI 183

QY 211 --GIDEPLHI-----KRRVVKPG--FHSPWKSAYIRQHRIDTNWRGELKSPKVL 258
Db 184 AAGVTIPDHQPCDLNVHREFLKLOFEGDIFERAADKSRYLADKTEKNWNPIMGSAVL 243

QY 259 KGHDDHVITCLQFCNRYVSGSDNTLKVMSAVTCKLRTLVLGHTGGVWSSQMRD--NII 316
Db 244 RGHEDHVITCMQIHDDVLVTGSDNTLKVWCDKGEVMTLVGHTGGVMTSQISQCGRYI 303

QY 317 ISGSTDRTLKVWNAETGECIHTLYGHTSTVRCMHLEKRVVSGSRDATLRVWDIETGOCL 376
Db 304 VSGSTDRTVVKWSTVDSGLLHTLQHTSTVRCMAMAGSILVTGSRDTRLRVWDVESGRHL 363

QY 377 HVLGMHVAARVCQYDGRVRYSGAYDFMVPDTPETCLHTLQHTNRVYSLOQFDGIH- 435
Db 364 ATLHGHAARVCQYDGRVRYSGAYDFMVPDTPETCLHTLQHTNRVYSLOQFDGIH- 423

QY 436 -VSSGLDTSIRVWDV---ETGNCIHTLTGHSQSLTSGMELKDNILVSGNADSTVKIMDIK 491
Db 424 IVCSGLDTSIRVWDVTRPEGOECVALLQHTSLTSGMELKDNILVSGNADSHVRVWDIH 483

QY 492 TGOCLOTQGPKNKHOSATVCLQ-FKNKEVITSSDGTGKVLMDLKTGEFIRNLVLTESGS 550
Db 484 EGTGVHMLSG---HRSATLSQWFGNRMVATSSDGTGKVLMDLKTGEFIRNLVLTESGS 540

QY 551 GGVVNRIRASTNKLVCAGVSRNGTEETKLLVLD 585
Db 541 GGCINRLCSTSTMCLACVSRNTEETKLLVLD 575

RESULT 2
CC4_CANAL          STANDARD;          PRT;          684 AA.
AC P53699;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 4.
GN CDC4.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGY126;
RA Shieh J.C., White A.M., Rosamond J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOKULATION. REQUIRED FOR HTAI-HTB1 LOCUS
CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X96763; CAA65538.1; -;  
InterPro; IPR001680; -;  
InterPro; IPR001810; -;  
Pfam; PF00646; F-box; 1;  
Pfam; PF00400; WD40; 7;  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS0181; FBOX; 1;  
DR PROSITE; PS00678; WD\_REPEATS\_1; 4;  
DR PROSITE; PS00082; WD\_REPEATS\_2; 4;  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1;  
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.  
FT DOMAIN 212 258 F-BOX.  
FT REPEAT 322 351 WD 1.  
FT REPEAT 363 391 WD 2.  
FT REPEAT 403 431 WD 3.  
FT REPEAT 442 468 WD 4 (POTENTIAL).  
FT REPEAT 478 506 WD 5.  
FT REPEAT 519 549 WD 6.  
FT REPEAT 561 589 WD 7.  
SQ SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;

Query Match 23.1%; Score 725.5; DB 1; Length 684;  
Best Local Similarity 30.8%; Pred. No. 2.2e-49;  
Matches 173; Conservative 102; Mismatches 181; Indels 105; Gaps 16;

QY 102 QORRITSVQ-----PPTGLQEWLKMFSQWSG-----PEKLLALDELIDSCPTQVQKHM 151  
Db 145 OGKANLDSIENVMSNEPTQSELVDLVTKLSGFLSEANQHLVFKLLQKTRTPT-LSTFN 203  
QY 152 QVIEPQFORDFISLLPKELALYVLSFLEPKDLQAQTCRVW-----RILAEDN 200  
Db 204 NLINSLARDILSNVPFEVTKILSYLDYKTLISVAQCKKMFIDINNPDTWIKLLKRDK 263  
QY 201 LLWREKCKEEGIDEPLHI-----KRRKVIKPGFIHSPWKSAYIRQHR 243  
Db 264 LITDDAVIKYELQYDQLLREWSTLPEINSQVLYKKRKIIV-----NRWMDPKPKPHRI 318  
QY 244 DTNWRREGELKSPKVLKGHDDHVITCLQFCNRYVSGSDNTLKVMSAVTCKLRTLVLGHT 303  
Db 319 S-----VSGHGNKVVTCLQHDDEKVVTVGDDKCSISYVSTQTGQLMKVLEGHE 365  
QY 304 GGVWSSQMRDNIIISGSDRTLKVWNAETGECIHTLYGHTSTVRCMHLEKRV-----356  
Db 366 GGVWALKYTGNTLVGTSTDRVTVNMNMTGCTHFRGHTSTIRCLDIHHPAVIGKNQDG 425  
QY 357 -----VSGSRDATLRWD-----IETGQ-----CLHVLGMHVAARVC 388  
Db 426 EDIVFPEYPLITGSRDHNHVMKLPVVVDDSDQYIETEGEFDNPLYIAVLSGHTQSVRS 485  
QY 389 VOYDGRVRYSGAYDFMVKVWD-PETETCLHTLQHTNRVYS--LOFDGIHVYSGSLDTSI 445  
Db 486 ISGVNIIISGSYSTVRVWDLDDGCHTHVQLGHLDRVYSTAIDFHSKTCFSGSMDNII 545  
QY 446 RVWDVETGNCIHTLTGHQSLSGMELKDNILVSGNADSTVKIMDIKTQCLOTQGPKNKH 505  
Db 546 NVWNFETGELKVLVGHASLVGLLDVLDVLSAAADATLRWDKATGELSKLKG---H 602  
QY 506 QSATVCLQFNKFVITSSDDGTGKVLMDLKTGEFIRNLVLTESGGSGGVVWRIRASTNKL 565  
Db 603 GAAITCFEHDGLRVVSGSEK-MLKLNWVNEKGFTEARDLLESDVTGS----IQVREDYKRCV 657  
QY 566 CAVGS--RNGTEETKLLVLD 584  
Db 658 AAVQRIINDEGETFIELD 678

Db 118 SFRILSFLDARSQCQAQVSKHWKELADDDVVIWHRMCEQHIHNRKCKGWLPLLRNTL 171  
Qy 210 -----EGID---EPLHIKRRKV-----IKP----- 236  
Db 178 YAAKASIQKRYELTRKRGVDOAHESPPVKAKLDDYPTSSNEETISSVKKPPSPNSDSKFF 237  
Qy 227 -GFTHSPKASYIROHIDTNWRGELKSPKVLKGHDDHVITCLQPCGNRIVSGSDNTL 285  
Db 238 LPFTKRWPKVEYAEKRCVECNWRHGRCRQ-VVLSGHSDBGVM-CLQLVRNLTASGSDATI 295  
Qy 286 KWSAVTGKCLRTLVLGHTGGWSSMRDNIIISGSTDRTLKVNAAETGECIHTLYGHTST 345  
Db 296 RLWNLATFQQVALLEGHSSGVTCLOFQDCKLISGSMDKTIRIWNRYTSECISLARGHTDS 355  
Qy 346 VRCMLHKEKRVSGSRDATLRVMDIETGQCLHLVIMGHVAAVRCVQY--DGRRVVSGAYDF 403  
Db 356 VLCLTFDSTLLVSGSADCTVKLWHFSGGKRI-TRLGHGTGPNVSVRIIRDGLVLSGSDDS 414  
Qy 404 MVKWDDEPETCLHTLQCHTRNVYSLQFDGIHVHVSGLDTSIRWVDVETGNCIHTLTGHQ 463  
Db 415 TIKIWSLETNCLHTGHTSAHIGPVOSIALADSRFCSLQDGIKQWIDIEKKKCVHTLFGH- 473  
Qy 464 SLTSGMEL-KDNI-LVSGNADSVTKWIDIKTGQCLQCTLOGPNKHQSAYTCLQPNKKNFVIT 521  
Db 474 -IEGWEIAADHLRLISGAHDGVVKWVE--ACEVHTLK---NHSEPVTVSVALGDCVVS 527  
Qy 522 SSDDGTGTVKLW 531  
Db 528 GSEDGKIYLW 537  
  
RESULT 4  
NT30\_YEAST STANDARD; PRT; 640 AA.  
AC P39014;  
CD 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE MET30 PROTEIN.  
GN MET30 OR YIL046W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=X2180-1A;  
RX MEDLINE=96069360; PubMed=8524217;  
RA Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,  
RA Surdin-Kerjan V.,  
RT "Met30p, a yeast transcriptional inhibitor that responds to S-  
RT adenosylmethionine, is an essential protein with WD40 repeats.",  
RL Mol. Cell. Biol. 15:6526-6534(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
RA Walsh S.V., Whitehead S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- GENES EXPRESSION.  
CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.  
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT  
CC PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collabora







Db	316	LIHNEAVLHLRNSGLMWTCSDKRSIAVMDMASATDITLRRVLVGHRAAVNVVDFDDKY	375
Qy	396	VVSGAYDFMKVVKWDPETETCLHTLQGHTRNVYSLQFDGHHVSGSDTSTRVMDVTGNC	455
Db	376	IVSAGDRTIKVWSTSCFEVTRLNGHKRIACLYRDLRVSSGSDNTRLMDIECGAC	435
Qy	456	IHTLTGHQSTTSGMELKDNLTIVSGNADSTVKIWDIKTG-----OCLTQLGGPNKHQ	506
Db	436	LRVLEGHEELVRCIRFDNKRIVSGAYDGKIKYWDLQAALDPRAPASTLCLRTLV--EHS	492
Qy	507	SAVTCLOFNKNFVITSSDDGTVKLWD	532
Db	493	GRVFRQLQDFEQIISSHDDTLIID	518

RESULT	7		
POP2_SCHPO			
ID	POP2_SCHPO	STANDARD;	PRT; 703 AA.
AC	Q14170;		
DT	15-JUL-1998	(Rel. 36, Created)	
DT	13-JUL-1998	(Rel. 36, Last sequence update)	
DT	01-OCT-2000	(Rel. 40, Last annotation update)	
DE	WD-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).		
GN	POP2 OR SUDI OR SPAC4D7.03.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OC	Schizosaccharomycetes.		
OX	NCBI_TaxID:4896;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RA	Wolf D.A., Jackson P.K.;		
RL	Submitted (DEC-1997) to the EMBL/GenBank/DBSJ databases.		
[2]			
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=98318628; PubMed=9653157;		
RA	Jallepalli P.V., Tien D., Kelly T.J.;		
RT	"sudi" targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1		
RT	proteins for degradation and stops unwanted diploidization in fission		
RT	yeast.;"		
RL	Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).		
RN	[3]		

SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA Gentles S., Churcher C.M., Barrrell B.G., Rajandream M.A., Wood V.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PUTATIVE COMPONENT OF E3 (PROTEIN-UBIQUITIN LIGASE)  
CC COMPLEX; TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND  
CC RUM1 PROTEINS FOR DEGRADATION.  
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL: AF038867; AAB95480.1; -
DR	EMBL: AF064515; AAC39496.1; -
DR	EMBL: Z98602; CAB11275.1; -
DR	InterPro: IPR001680; -
DR	InterPro: IPR001810; -
DR	Pfam: PF00646; F-box.1.
DR	Pfam: PF00400; WD40.6.
DR	PRINTS: PR00320; GPROTEINBRPT.
DR	PROSITE: PS50181; FBOX.1.
DR	PROSITE: PS00678; WD_REPEATS_1; 3.
DR	PROSITE: PS50082; WD_REPEATS_2; 6.
DR	PROSITE: PS50294; WD_REPEATS_REGION.1.

[illegible]

Query Match	20.4%	Score 641;	DB 1;	Length 703;
Best Local Similarity	27.3%;	Pred. No. 1e-42;		

Qy	9	LNHGLVPVDLKAKEPL-----PHOTWNKIFSIIAQ-----GLP-----F	45
Db	85	LKHSCTSFRLKFNENSYKCLDHSPDPSGDFESILVPQRFNFXSHSLPPKIIIS	144
Qy	46	CRRMRKRLDHGSEVRFSGLGKPKCKYSEVTSITGLVPCSATPTTFCDLRAAANGCQQR	105
Db	145	IDRNRKIKLONSISSGDNPPPKYDTSNTVSPGSKPISE-----DLEDLQDSIVQ-	197
Qy	106	RITSVQPTGLQEWLKFQWSGPEKLLALDELIDSCPTQVAKHMQVIEPQRFOSIL	165
Db	198	-TEEDLEPGIQSY-----AFFQLLRSCNRQSMRLLENECEPLLKKDILSN	241
Qy	166	LPKELALIVLSFLEPKDLQAQTCRYW-RILAEADNLLWREKCKEGI---DEPLHIKRR	221
Db	242	LPFIVOSIILNLDIHSLFSLRVSPPTNRILOVHTSYWKHMFSLFCFQINENDWKYAP	301
Qy	222	KVIKPGITHSPWKS-----AYIRQHRIDTNRRGELKSPKVLKGH-----DDHVITCL	269
Db	302	NLNRPEPLHNDQISDDYFPEIFKRHFLN-----KRWLFPSIPSHLSFPPIHVNFMTSL	357
Qy	270	QFCNCRITVSGDDNTLKWSAVTKGKCLRTLVLGHTGGVWVSSQMRDNIISGSTDRTLKVN	329
Db	358	LLHKDRITITSGSGTQIHNAITGVLEARLEGHKEGVAVKIHENTLVSGSIDKTVRVVN	417
Qy	330	AETCEGITHLYGHTSTVRCHL-----HEKRVVSGSRDATLRVWDI---370	
Db	418	IEKAKCHIFRFGHISIRLCLEILVPSRLIRHGEIVVEPDOPYIVSGSRDHTLRWKLPN	477
Qy	371	-----ETQCQLHVLHGVAAVRCVQYDGRVVSGAYDFMVKWVPDPEE	413
Db	478	TDPPLPDNTNSIDRWEKNPYFVHTLIGHTSDTSVRTISGYGDIILVSGSYDSSIRIWRVSTG	537
Qy	414	TCLHTLOGHNRVYSLOFDGIH--VYSGSLDTSIRVWDVETGNCIHTLTHGQSLTSGMEL	471
Db	538	ECYLHGRHSRLYSVLYEPERNICISGSMKDSIRVWDLSTGCKVYLEGHDAFVTLN	597
Qy	472	KDNLIVSGNADSTVKIWDIKTCLOTLQCPNKHQSAVTCLOENKNFVITSSDDGVFKVL	531
Db	598	FQNLISGSADSIIRIWDNLNTPPLVLPNSNGYSISFV-----SDEHKIISGNDGSVKLW	653
Qy	532	DLKTGEFIRNLVLESGSGGVVNRIRASNTKLVCVAGSRNGTEETKLLVLDF	594
Db	654	DVRTKLLRELLT-----DLTKTHWVDFDAMRCVAAV---QRDOAYLEVIN	698

RESULT	8
YSSL_CAEEL	
ID	YSSL_CAEEL
STANDARD;	PRT; 701 AA.
AC	Q09930;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	DE HYPOTHETICAL 80.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN K10B2.1 IN
DE	CHROMOSOME II.
GN	K10B2.1.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Peloderinae; Caenorhabditis.



```
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS00181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 6.
DR PROSITE; PS0082; WD_REPEATS_2; 7.
DR PROSITE; PS0082; WD_REPEATS_2; 7.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
FT DOMAIN 190 228 F-BOX.
FT REPEAT 301 338 WD 1.
FT REPEAT 341 378 WD 2.
FT REPEAT 381 418 WD 3.
FT REPEAT 424 461 WD 4.
FT REPEAT 464 503 WD 5.
FT REPEAT 505 541 WD 6.
FT REPEAT 553 590 WD 7.
FT VARSPLIC 17 52 MISSING (IN ISOFORM 2).
SQ SEQUENCE 605 AA; 68866 MW; 4C67F3B7E400FD37 CRC64;

Query Match 20.3%; Score 638; DB 1; Length 605;
Best Local Similarity 30.8%; Pred. No. 1.4e-42;
Matches 166; Conservative 92; Mismatches 209; Indels 72; Gaps 15;

Qy 56 HGVSRFSGLKKPKV-----SEVSTTGL-----VPCSA-----TPTTFGLRAA 97
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 NSSRDCCNGEPPRKIIPEKNSLRQTVNSCARCLNQETVCLASTAMKTCNCAKTCLA 112
Qy 98 NGQG-----QQRRTTSVQPTGLQEWLKFQSWSGPEKLLALDELIDSCPTQVKHMQ 152
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
113 NGTSMVVPKOKLSAYEKEKEL--CVKYFEQMSQDFVEHLISQMHYQGHINS 170
Qy 153 VIEPQFQDFTSLP-----KELALVLSFLPKDLLQAAQCYRWIRLAENLLWRECK 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 YLKPMLQDFITLALPARGLDHIAENILSYLDAKSLCAELVCKEYRVTSQMLMKLIE 230
Qy 209 E-----EGIDEPLHKRRK-----VIRPGFTSPKWSAY-----IRQ--HRIDTN 246
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 RMVRTDSLWRLAE-----RRGWGYLFKNKPPDGNAPNSFYRALYPKIIQDIETIESN 285
Qy 247 WRGELKSPKV-LKGDDHVTCLQFCGNRIVSGDDNTLVKWSAVTGKCLRTLVGHGTG 305
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
286 WRCGRHSRLQIRHCRSETSGVYVQDQKIVSGRLDNTIKIWDKNTLECKRIITGHTGS 345
Qy 306 VWSSQMRDNIISGSTDTLKWNNATGECITLYGHTSTVRCMHLHEKRVVSGSRDATL 365
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 VLCLQYDERVIITSSDSTVRVMDYNTGEMTLTHHCEAVLHLRFNMGVMVTCSDRSI 405
Qy 366 RVWDETGOCL---HVLGMHVAARVCVYDGRVRYVSGAYDFMKVMDPETETCLHTLQGH 422
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 AVWDMASPTDILRLVLRVGHRAVNVDFDKYIVSASGDRITKVNWTSTCEFTVLNGH 465
Qy 423 TNRVYSLQFDGHHVYSGSLDTSIRVMDVETGNCIHTLTHGSHSLGSMELKDNILVSGNAD 482
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
466 KRGIACLOYRDLRVYSGSDNTIRLWDIECGACLRVLSGHEELVRCIRFDNKRIVSGAYD 525
Qy 483 STVKTIWDITKG-----OCLQTLQGNKHQSVAVTCQLQKNKVIITSSDGTGVKLD 532
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
536 GKIKVMDLVAALDPRAPAGTCLRLTLV---BHSGRVRLQFDEFOIVSSSHDDTLIWD 581

RESULT 10
TRCB_XENLA
ID TRCB_XENLA STANDARD; PRT; 518 AA.
AC Q91854; P70038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
GN FBXW1 OR BTRCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
```

```

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93330289; PubMed=8393141;
RX Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
RA "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
RT anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
FT with beta-transducin repeats.";
RL Mol. Cell. Biol. 13:4953-4966(1993).
[2]
RN SEQUENCE OF 302-518 FROM N.A.
RX MEDLINE=97109804; PubMed=8952061;
RA Hudson J.W., Alarcon V.B., Elinson R.P.;
RT "Identification of new localized RNAs in the Xenopus oocyte by
FT differential display PCR.";
RL Dev. Genet. 19:190-198(1996).
CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
CC TADPOLE EMBRYO.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M98268; AAA02810.1; -
EMBL; U63921; AAB49671.1; -
EMBL; U63922; AAB49672.1; -
InterPro; IPR001680; -
InterPro; IPR001810; -
Pfam; PF00646; F-box; 1.
Pfam; PF00400; WD40; 7.
PRINTS; PR00320; GPROTEINBRPT.
PROSITE; PS00181; FBOX; 1.
PROSITE; PS00678; WD_REPEATS_1; 6.
PROSITE; PS0082; WD_REPEATS_2; 7.
PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Ubiquitin conjugation; Repeat; WD repeat.
DOMAIN 119 157 F-BOX.
FT REPEAT 230 258 WD 1.
FT REPEAT 270 298 WD 2.
FT REPEAT 310 338 WD 3.
FT REPEAT 353 381 WD 4.
FT REPEAT 393 421 WD 5.
FT REPEAT 433 461 WD 6.
FT REPEAT 482 510 WD 7.
FT CONFLICT 302 304 GEM -> EFR (IN REF. 2).
FT CONFLICT 516 518 GLA -> AAH (IN REF. 2).
SQ SEQUENCE 518 AA; 59507 MW; 2A52EC19028127F3 CRC64;

Query Match 20.2%; Score 634.5; DB 1; Length 518;
Best Local Similarity 31.1%; Pred. No. 2.2e-42;
Matches 161; Conservative 90; Mismatches 214; Indels 53; Gaps 12;

Qy 60 VRSFSLGKPKCKVSEYTTTGLVPCSAIPT---TFGLRAANGQSQ---RRRTTSVQPP 113
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MEGFSCSLQPTTASERDCNRDEPPRKIIITEKNTLRQTKLANGTSMIVPKOKLSANYE 60
Qy 114 TGLQWLKMFQSWSGPEKLLALDELIDSCPTQVKHMQVIEPQFQDFTSLP----KE 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 KEKELCVKYFEQMSQDQVEFVEHLISRMCHYQGHINTYLPKMLQDFITLALPARGLDH 120
```



DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE CELL DIVISION CONTROL PROTEIN 4.  
GN CDCA4 OR YFL009W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID:4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88011240; PubMed=3309335;  
RA Yochem J., Byers B.;  
RT "Structural comparison of the yeast cell division cycle gene CDC4 and  
RT a related pseudogene.";  
RL J. Mol. Biol. 195:233-245(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=93400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae.";  
RL Nat. Genet. 10:261-268(1995).  
RN [3]  
RP SEQUENCE OF 1-579 FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Barrrell B.G., Churcher C., Rajandream M.A.;  
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA  
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE  
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD  
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND  
CC VARIOUS ASPECTS OF SPOULATION. REQUIRED FOR HTA1-HTB1 LOCUS  
CC TRANSCRIPTION ACTIVATION.  
CC -!- SIMILARITY: CONTAINS 1 F-BOX DOMAIN  
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X05625; CAA29113.1; -  
DR EMBL; D50617; BAA09229.1; -  
DR EMBL; D31600; BAA06495.1; -  
DR EMBL; Z46255; CAA86341.1; -  
DR PIR; A26867; A26867.  
DR SGD; S0001885; CDC4.  
DR InterPro; IPR001680; -  
DR InterPro; IPR001810; -  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00400; WD40; 6.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS50181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS\_1; 4.  
DR PROSITE; PS50082; WD\_REPEATS\_2; 5.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.  
FT DOMAIN 272 319 F-BOX.  
FT REPEAT 380 408 WD 1.  
FT REPEAT 420 449 WD 2.  
FT REPEAT 461 493 WD 3.  
FT REPEAT 528 556 WD 4.  
FT REPEAT 568 598 WD 5.  
FT REPEAT 630 658 WD 6.  
FT REPEAT 669 698 WD 7.  
FT CONFLICT 460 460 K -> E (IN REF. 1).  
SQ SEQUENCE 779 AA; . 86089 MW; 0348F2F8FA78F3BC CRC64;

Query Match 19.7%; Score 618.5; DB 1; Length 779;  
Best Local Similarity 29.6%; Pred. No. 6.9e-41;  
Matches 152; Conservative 105; Mismatches 176; Indels 81; Gaps 16;  
Qy 135 LDELIDSEPTQVHKMQVIEPQFORDFISLLPKELALYVLSFLEPKDLLAAQAQTCRW- 193  
Db 247 LFRLVANMDRSELSDLGLTIKLDNLKRDILTSLPFEISLKFNYLQFEDIIISLGSQNN 306  
Qy 194 RILAEDNLLWREKCKEEDIDP-----LHKRRKVIKPGFIHSPKSAVIHQHRTDWNR 248  
Db 307 KIIRKSTSLWKKLLISENFVSPKGFNSNLKSLQKPKLSQQDRURLSFLFNIFILKNM- 365  
Qy 249 RGEKSPK-----VLKGDHDIITCLOFCGNIIVSGDDNTLKVWSAVTQKCLRTLVGH 302  
Db 366 ----YNPKEVFPQRTTLRGHMTSVITCLOFEDNYITGADDKMIRVYDSINKFLQLSGH 421  
Qy 303 TGGVMS-SQMRDNIISGTDRTLKVNAETGECIHTLYGHTSTVRCMHLE-----KRVV 357  
Db 422 DGGWALAYAHGGLVSGTDRTVRVWDIKKGCCTHVEKGNSTVRCLDIVYEKKIKYIV 481  
Qy 358 SGRDATLRVWDI-----ETGQCLHVLGMGHVAARVCVOYDGR 394  
Db 482 TCSRDNLTLRVWKLKPKRESSVPDHGEHDYPLVFHTPEENPYFVGVLRGHMASVRYTSGHN 541  
Qy 395 RVSGAYDFNVKVPDPTETCLHTLQHTNRVYSLOFDGIH---VVSGLDTSIRVWDV 450  
Db 542 IVSGSYDNTLIVDVAQMKCLYILSGTDRIYTYD--HERKRCISASMDTTIRIMDL 599  
Qy 451 E----TGNC-----IHTLTGHQSLTSGMELKDNILVSGNADSTVKIWDI 490  
Db 600 ENIWNNGECVATNSASPCAKILGAMYTLQHTALVGLRLSDKFLVSAADGSRGWD- 658  
Qy 491 KTGQCLQTLQGNKHQSVAVTCLQFNKNFVTSDDGTGTVKMLDKTGEFIRNLVLESQGS 550  
Db 659 -ANDYSRKFYSYHHTNLSAITFYVSDNIVSGSEN-QFNIYNLRSGKLHVANILKDA--- 713  
Qy 551 GGVWRIIRASNTKLVCAVGSNRNGTEETKLLVDF 584  
Db 714 -DQIWSVNFKGKTLVAHV-EKDG--QSFLEILDF 743

## RESULT 13

HET1\_PODAN  
ID HET1\_PODAN STANDARD; PRT; 1356 AA.  
AC Q00808;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE VEGETATIVE INCOMPATIBILITY PROTEIN HET-E-1.  
GN HET-E1.  
OS Podospora anserina.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Podospora.  
OX NCBI\_TaxID=5145;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96009891; PubMed=7557402;  
RA Saube S., Turcq B., Begueret J.;  
RT "A gene responsible for vegetative incompatibility in the fungus  
RT Podospora anserina encodes a protein with a GTP-binding motif and G  
RT beta homologous domain.";  
RL Gene 162:135-139(1995).  
CC -!- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH  
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,  
CC HET-C.  
CC  
CC -!- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way





QY 228 -----FIHSP-----WKSAYIRQHRIDTNRNGELKSPKYLKGHDHVTITCLOFCQNR 275  
Db 249 MAAEASKAVTQPKTRSMKAVYDRWQVSYNNKNSRYKL-SVLKGHENG-V-TCLQLDDNI 306  
QY 276 IVSGDDNTLKVSAVTKCLRTLVGHGTGGVWSSQMRDNIISGSTDRTLKVNNAETGEC 335  
Db 307 LATGSYDTTIKIWNITETECIRTLVGHGTAGRALQFDOSKLISGLDHTIKVNNWHTGEC 366  
QY 336 IHTLYGHTSTVRCHLHKKRVRVSGSRDTRLRWDTLETQCLHVLGMHVAARVCOYD--G 393  
Db 367 LSTFAARTDSVLSVHFHGLHLLASGSGDKTVKIFDFNSKE-TYCLKGSDWNSTHVDIKS 425  
QY 394 RRVVSGAYDFMVKVWDPETETCLHTLQGTNRVYSLQF----- 431  
Db 426 RIVFSASDDTIKLDLDTQRVIRYEGHVQVQVILPPEYEPDEEVLNGASQDNQDA 485  
QY 432 -----DGIH----- 435  
Db 486 MSVSSGGSGSPMSHAQIERAGSPGSHSSHNLLPSSLPSGDEVDYRLHLYGSFAVADESRP 545  
QY 436 -----VVSGLDTSRVWDVETGNCIHTLTH-----QSLTSGMELKDN-LVSGNADSTV 485  
Db 546 LPRIYFMTGGDLSTWRLWDSATGRCLRTLFGHLEGVWSLAG-----DTRIVISGANDGMV 600  
QY 486 KTWDLTKTGCLQTLQGNPKHQSVAVTCLOFNKNFVITSSDDGTGTVKLWDLK 534  
Db 601 KTWPRSGKCDATYTG---HCGPITCVGLSDSLMASGSEDGTIRUHSFK 646  
  
RESULT 15  
KMBH\_DICDI  
ID KMBH\_DICDI STANDARD; PRT; 732 AA.  
AC P90648;  
CD 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE MYOSIN HEAVY CHAIN KINASE B (EC 2.7.1.129) (MCKB B).  
GN MCKB OR MCKB.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AX3;  
RA MEDLINE=97277316; PubMed=9115238;  
RX Clancy C.E., Mendoza M.G., Naismith T.V., Kolman M.F., Egelhoff T.T.;  
RT "Identification of a protein kinase from Dictyostelium with homology  
to the novel catalytic domain of myosin heavy chain kinase A.";  
RL J. Biol. Chem. 272:11812-11815(1997).  
CC -!- FUNCTION: PHOSPHORYLATES THREONINE IN THE C-TERMINAL TAIL REGION  
OF MYOSIN II HEAVY CHAIN. THIS PHOSPHORYLATION IS CRITICAL IN  
REGULATING THE ASSEMBLY AND DISASSEMBLY OF MYOSIN II FILAMENT.  
CC -!- CATALYTIC ACTIVITY: ATP + [MYOSIN HEAVY-CHAIN] = ADP +  
[MYOSIN HEAVY-CHAIN] PHOSPHATE.  
CC -!- DOMAIN: CONSISTS OF AN N-TERMINAL DOMAIN, A CENTRAL NONREPETITIVE  
CATALYTIC DOMAIN, AND A C-TERMINAL DOMAIN WITH SEVEN WD REPEATS.  
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
CC -!- SIMILARITY: BELONGS TO THE MCKB / EF-2 PROTEIN KINASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U09046; AB50136.1; -.  
CC DictyDb; DD01087; mhkb.  
CC InterPro; IPR001680; -.  
CC Pfam; PF00400; WD40; 7.  
CC PRINTS; PR00320; GPROTEINBRPT.  
CC PROSITE; PS00678; WD\_REPEATS\_1; 5.

DR PROSITE; PS0082; WD\_REPEATS\_2; 6.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
KW WD repeat.  
FT NP\_BIND 298 303 ATP (POTENTIAL).  
FT DOMAIN 355 381 POLY-ASN.  
FT REPEAT 458 486 WD 1.  
FT REPEAT 500 528 WD 2.  
FT REPEAT 540 568 WD 3.  
FT REPEAT 580 608 WD 4.  
FT REPEAT 620 648 WD 5.  
FT REPEAT 660 688 WD 6.  
FT REPEAT 700 730 WD 7.  
SQ SEQUENCE 732 AA; 83167 MW; A7233C4BD56D4088 CRC64;  
  
Query Match 16.4%; Score 516; DB 1; Length 732;  
Best Local Similarity 34.0%; Pred. No. 7.2e-33;  
Matches 113; Conservative 72; Mismatches 111; Indels 36; Gaps 7;  
  
Qy 231 SPWKSAYTRQ--HRIDTNRRGELKS-----PKVL-----KGHDD 263  
Db 403 SPSRQLFVSDNGTNTLNKRSKSKSIDLEKPEILINNKKESINLETKLIETIKGY-- 460  
Qy 264 HVITCLOFCGSRIVSGSDNTLKV--WSAVTCCLRTLVLGHGTGGVWSSQMRDNIISGST 321  
Db 461 HVTSHLCICDNLFFTGCSDNSIRVDYKSONNECVQTLKGHEGPVESICYNQYLFSGSS 520  
Qy 322 DRTLKVNNAETGECIHTLYGHTSTVRCMHLHKKRVVSGSRDTRLRWDTETGQCLRVLMG 381  
Db 521 DHSIKVWDLKKLRCIFTLEGHDKPVHTVLLNDKYLFGSSDKTIKVDLKTLECKYTLES 580  
Qy 382 HVAAVRCVOYDGRVSVGAYDFMVKVWDPETETCLHTLQGTNRVYSLQFDGIHVVSSGL 441  
Db 581 HARAVKTLICISQYLFSGSDNKTIKVWDLKTRFCNTYTLKGHTKWTTCILGTNLVSGSY 640  
Qy 442 DTSIRVWDVETGNCIHTLTGHSLTSGMELKDNILVSGNADSTVKIMDKITGQCLTLOG 501  
Db 641 DKTIRVWNLKSLKESATLRGHDWRVVEHVMVICDKLLFTASDDNTIKIWDLLETURCNTLEG 700  
Qy 502 PNKHQSVAVTCLOF--NKNFVITSSDDGTGTVKWLW 531  
Db 701 ---HNATVQCLAVWEDKKCVISCSHDSIRVW 729  
  
Search completed: September 25, 2001, 14:45:03  
Job time: 208 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2001, 14:41:10 ; Search time 22:58 seconds  
(without alignments)  
3451.181 Million cell updates/sec

Title: US-09-328-877A-8

Perfect score: 3143

Sequence: 1 MSKFGKPTLHGLVFPVLDLS.....SRNGTEETKLELVDFDVMK 589

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2889	91.9	553	4 Q9NUX6	Q9nux6 homo sapien
2	2062.5	65.6	1326	5 Q9VZF4	Q9vzf4 drosophila
3	1238	39.4	587	5 Q44083	Q44083 caenorhabdi
4	650	20.7	665	5 Q9GNH6	Q9gnh6 caenorhabdi
5	635	20.2	569	11 Q9Z159	Q9z159 mus musculu
6	635	20.2	569	11 Q9QU15	Q9qu15 mus musculu
7	630	20.0	569	11 Q9R1G7	Q9r1g7 mus musculu
8	617	19.6	510	5 Q44382	Q44382 drosophila
9	617	19.6	510	5 Q9VDE3	Q9vde3 drosophila
10	533	17.0	506	3 Q9P7V1	Q9p7v1 schizosacch
11	463.5	14.7	942	5 Q96611	Q96611 dictyosteli
12	461	14.7	334	4 Q9NWX7	Q9nwx7 homo sapien
13	439	14.0	410	13 Q9PTR5	Q9ptr5 gallus gall
14	436	13.9	410	6 Q9GL51	Q9gl51 sus scrofa
15	426	13.6	330	4 Q9NUL4	Q9nul4 homo sapien
16	410.5	13.1	317	10 Q9M2Z2	Q9m2z2 arabidopsis
17	392.5	12.5	333	10 Q9SY00	Q9sy00 arabidopsis
18	392.5	12.5	454	4 Q9HA09	Q9ha09 homo sapien
19	391.5	12.5	454	11 Q9QUH1	Q9quh1 mus musculu

20	389	12.4	391	11	O35592	mus musculu
21	388	12.3	411	5	O96698	drosophila
22	386.5	12.3	277	11	O9R2A6	mus musculu
23	378.5	12.0	502	3	O74855	schizosacch
24	366.5	11.7	594	4	Q9H073	homo sapien
25	366	11.6	515	10	Q9FGS2	arabidopsis
26	365.5	11.6	404	5	Q9NDC9	caenorhabdi
27	349.5	11.1	553	5	Q9VVI0	drosophila
28	348	11.1	579	5	O76734	dictyosteli
29	347	11.0	514	11	Q9WUC8	rattus norv
30	347	11.0	690	5	O61585	strongyloce
31	342.5	10.9	514	4	O43660	homo sapien
32	338	10.8	513	11	O55039	mus musculu
33	336	10.7	307	5	Q9V878	drosophila
34	336	10.7	777	3	Q9USN3	schizosacch
35	333.5	10.6	494	5	Q18295	caenorhabdi
36	332	10.6	476	13	O93531	xenopus lae
37	331	10.5	436	5	Q9VAK0	drosophila
38	329	10.5	514	11	O9E0D4	mus musculu
39	328	10.4	347	5	Q9VPL0	drosophila
40	328	10.4	482	5	Q9VYQ9	drosophila
41	328	10.4	522	4	O43445	homo sapien
42	324.5	10.3	651	3	Q9P7I3	schizosacch
43	324	10.3	514	4	Q9H9A1	homo sapien
44	323	10.3	485	4	Q9NVX2	homo sapien
45	323	10.3	520	4	O43864	homo sapien

#### ALIGNMENTS

RESULT 1

Q9NUX6

-ID Q9NUX6 PRELIMINARY; PRT; 553 AA.

AC Q9NUX6;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE CDNA FLJ11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10  
 DE PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK001933; BAA91986.1;  
 DR InterPro: IPR001680;  
 DR InterPro: IPR001810;  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PROSITE: PS0181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS; UNKNOWN\_5.  
 DR SMART: SM00256; FBOX; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 553 AA; 63280 MW; CA829C221986A3F2 CRC64;

Query Match 91.9%; Score 2889; DB 4; Length 553;

Best Local Similarity 99.8%; Pred. No. 2.8e-237;  
 Matches 540; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 49 RMRRKLDHGSEVRSFSGKKPKCKVSEYTSITGLVPCSATPTTFGLRAANGQGOORRIT 108

DB 13 KMRRKLDHGSEVRSFSGKKPKCKVSEYTSITGLVPCSATPTTFGLRAANGQGOORRIT 72

QY 109 SVOPPTGLQEWLKMFSQWSGPKLLALDELIDSCPTQVKHMMQVIEPQFQDFISLLPK 168  
Db 73 SVOPPTGLQEWLKMFSQWSGPKLLALDELIDSCPTQVKHMMQVIEPQFQDFISLLPK 132  
QY 169 ELALYVLSFLEPKDLLQAAQTCRYWRIIAEDNLLWRECKKEBGIDEPHLIKRRKVIKPGF 228  
Db 133 ELALYVLSFLEPKDLLQAAQTCRYWRIIAEDNLLWRECKKEBGIDEPHLIKRRKVIKPGF 192  
QY 229 IHSPKSAVIROHRIDTNNRRELKSPVKLGKDDHVTICLQFCGNRIVSGDDNTLKW 288  
Db 193 IHSPKSAVIROHRIDTNNRRELKSPVKLGKDDHVTICLQFCGNRIVSGDDNTLKW 252  
QY 289 SAVTKCLRTLVGHRGVMSSQMRNIIISGSDRTLKWNAETGECIHTLGHSTVRC 348  
Db 253 SAVTKCLRTLVGHRGVMSSQMRNIIISGSDRTLKWNAETGECIHTLGHSTVRC 312  
QY 349 MHLHKRVVSGSRDATLRVMDIETGCLHVLGMHVAARVCQYDGRVVGSGAYDFMVKW 408  
Db 313 MHLHKRVVSGSRDATLRVMDIETGCLHVLGMHVAARVCQYDGRVVGSGAYDFMVKW 372  
QY 409 DPETETCLHTLQGHNRVYSLQFDGIHVVGSLDTSIRVMDVETGNCIHTLGHQSLTSG 468  
Db 373 DPETETCLHTLQGHNRVYSLQFDGIHVVGSLDTSIRVMDVETGNCIHTLGHQSLTSG 432  
QY 469 MELKONILVSGNADSTVKIWDIKTGOCLOTLOGPNKHOSAVTCLQFNKNFVITSSDDGTV 528  
Db 433 MELKONILVSGNADSTVKIWDIKTGOCLOTLOGPNKHOSAVTCLQFNKNFVITSSDDGTV 492  
QY 529 KLWDLKTGEFIRNLVLTESGGGVVWRIRASNTKLCAVGSRNGTEETKLVLVDFDVM 588  
Db 493 KLWDLKTGEFIRNLVLTESGGGVVWRIRASNTKLCAVGSRNGTEETKLVLVDFDVM 552  
QY 589 K 589  
Db 553 K 553  
RESULT 2  
QYVZF4 PRELIMINARY; PRT: 1326 AA.  
AC QYVZF4;  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)  
DE CG15010 PROTEIN.  
GN CG15010.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
RX NCBI\_TaxID:7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003480; AAC22246.1; .  
DR Flybase; FBgn0035516; CG15010.  
DR InterPro; IPR001680; .  
DR InterPro; IPR001810; .  
DR Pfam; PF00400; WD40; 7.  
DR Pfam; PF00646; F-box; 1.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR PROSITE; PS00181; FBOX; 1.  
DR PROSITE; PS00678; WD\_REPEATS; 5.  
KW Repeat; WD repeat.  
SQ SEQUENCE 1326 AA; 141360 MW; 3F42C873CFA3027F CRC64;  
Query Match 65.6%; Score 2062.5; DB 5; Length 1326;  
Best Local Similarity 66.8%; Pred. No. 1.8e-166;  
Matches 405; Conservative 54; Mismatches 98; Indels 49; Gaps 10;  
QY 12 GLVPVDLKAKEPLPHQTVMKI-----FSTSIQAQ-----LFCRRMR 52  
Db 733 GTEYSLTNPSSLMGSHQRKKEGRLLDGGDYSVTISSSGEVGGPGSVSNCRKRI-- 790  
QY 53 KLDHGSEVRSFS-LCKKPKCVSEYVSTTCLVPCSATPTFGDLRAANGOGQORRITSV- 110  
Db 791 AYDFASTPRSSQHLG--PTAVLSVTPSSHLT--SSTP-----GSALGRRTPRSVP 836  
QY 111 ---QPPTGLQEWLKMFSQWSGPKLLALDELIDSCPTQVKHMMQVIEPQFQDFISLLP 167  
Db 837 SRDNPPELQHWLAQFQWVSHVERLLALDLHDGCDPSQVRHMMKVIEPQFQDFISLLP 896  
QY 168 KELALYVLSFLEPKDLLQAAQTCRYWRIIAEDNLLWRECKKEGI-----DEPLHKRR 221  
Db 897 RELALFVLSYLEPKDLLRAAQTCRSWRFLCDDNLLWKECKRKAQILAEPRSDP---KRG 953  
QY 222 KVIKPGFTHSPWKSAYIRQHRIDTNWRGELKSPVKLGKDDHVTICLQFCGNRIVSGSD 281  
Db 954 RDGNPPPIASPKAAAYMRQHIIEMNWRSPRPKPKVKGKDDHVTICLQFCGNRIVSGSD 1013  
QY 282 DNTLKWVSAVTKCGLRTLVGHTGGVWSSQMRNIIISGSDRTLKWNAETGECIHTLYG 341  
Db 1014 DNTLKWVSAVTKCGLRTLVGHTGGVWSSQMRNIIISGSDRTLKWMDMSGCVHTLQG 1073  
QY 342 HTSTVRCMHLHKRVVSGSRDATLRVMDIETGCLHVLGMHVAARVCQYDGRVVGSGAY 401  
Db 1074 HTSTVRCMHLHKRVVSGSRDATLRVMDIETGCLHVLGMHVAARVCQYDGRVVGSGAY 1133  
QY 402 DPMVKWDPETETCLHTLQGHNRVYSLQFDGIHVVGSLDTSIRVMDVETGNCIHTLFG 461  
Db 1134 DPMVKWDPETETCLHTLQGHNRVYSLQFDGIHVVGSLDTSIRVMDVETGNCIHTLFG 1193  
QY 462 HOSLTSGMELKDNILVSGNADSTVKIWDIKTGOCLOTLOGPNKHOSAVTCLQFNKNFVIT 521  
Db 1194 HOSLTSGMELKDNILVSGNADSTVKIWDIKTGOCLOTLOGPNKHOSAVTCLQFNKNFVIT 1253



```

      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 420 VLEGHLELVRCIRDEKRVISGAVDGKIKVWDLQAALDPRALSSICLSLV---QHTGR 476

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 509 VTCLOFNKNFVITSSDDGTVKLWD 532
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 477 VFLQFDQFQIVSSSHDDTILWD 500

RESULT 5
Q92159 ID Q92159 PRELIMINARY; PRT; 569 AA.
AC Q92159;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99145465; PubMed=9990853;
RA Spencer E., Jiang J., Chen Z.J.;
RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
RT Slimb/beta-TrCP."
RL Genes Dev. 13:284-294(1999).
DR EMBL; AF112979; AAD04181.1; -.
DR InterPro; IPR001680; -.
DR InterPro; IPR001810; -.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS50181; FBOX; 1.
DR SMART; SM00320; WD40; 1.
DR SMART; SM00678; WD_REPEATS; UNKNOWN_6.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

      | | | | | : : : : | | | | | : | | | | : | | | | :
Query Match 20.2%; Score 635; DB 11; Length 569;
Best Local Similarity 30.3%; Pred. No. 1.3e-45;
Matches 164; Conservative 92; Mismatches 214; Indels 72; Gaps 14;

Qy 53 KLDHGEVRSFSLGKKPKCV-----SEYTSITGLV-----PCSATPTTFGDL 94
      | | | | | : | | | | : | | | | : | | | | :
Db 14 KFMNSSEDCNCGEPKPKIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTCNCVAKA 73

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 95 RAANGOG-----QRRRTTSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKH 149
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 74 KLANGTSSMIVPKQKLSASYEKEKEL--CVKYFEQWSESDQVEFVHLISQMCHYQGHG 131

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 150 MMQVTEPQFQDFISLSP-----KELALYVLSFLEPKDLLQAAQTCRYWRIIAEDNLLWRE 205
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 132 INSYLKPMLQDFITALPARGLDHIAENILSYLDKASLCAELVCKEYRVTSQGLMKWK 191

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 206 KCKE-----EGIDEPLHKRRK-----VIRPGFIHSPKSAI-----IRQ--HRI 243
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 192 LIERMVRTDSLWRGLAE-----RRGWGQYLFKNKPPDENAPPNSFYALYPKIIQDIETI 246

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 244 DTNWRGELKSPKV-LKHGDHVVITCLQFCGNRIVSGDDNTLVKWSAVTGKCLRTLVLGH 302
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 247 ESNWRCGRHSLQRICHRSETSGVYVCLQYDQKIVSGLRDNTIKIWDKSTLECKRIITGH 306

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 303 TGGVWSSQMRNIIISGSDRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 362
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 307 TGSVLCLQYGERVIITGSSDSTRVWMDVWAGEMLNTLHHCEAVLHFRNMGMMVTCSDK 366

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 363 ATLRRVWDIETGQCL---HVLGMHVAARCVQYDGRVRYSGAYDFWVKWMDPETETCLHTL 419
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 367 RSIADWDMASPTDITLRRVLVGHRAAVNVDFDDKIYVSGSDRTIKVWNTSTCEFEVRTL 426

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 420 QGHTNRVYSLOFDGIHVVGSLDTSIRVMDVETGNCIHTLTGHOSLTSGMELKONILVSG 479
```

```

      | | | | | : | | | | : | | | | : | | | | :
Db 427 NGRKRGACIACLRDLRVVSSGSDMTIRLWDIECCACLRVLEGHEELVRCIRFDNKRIYVG 486

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 480 NADSTVKIWDIKTG-----OCLQTLQGNKHKQSAVTCLQFNKNFVITSSDDGTVKL 530
      | | | | | : | | | | : | | | | : | | | | :
Db 487 AYDGKIKVWDLMAALDPRAPAGTLCRLTLV---EHSGRVFLQDFQIVSSSHDDTILI 543

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 531 WD 532
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 544 WD 545

RESULT 6
Q9QUI5 ID Q9QUI5 PRELIMINARY; PRT; 569 AA.
AC Q9QUI5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE UBIQUITIN LIGASE FWD1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99199275; PubMed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-I.;
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul 1/F-box protein FWD1."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase."
RL Nature 396:590-594(1998).
DR EMBL; AF081887; AAD17755.1; -.
DR EMBL; AF099932; AAD08701.1; -.
DR InterPro; IPR001680; -.
DR InterPro; IPR001810; -.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_6.
DR SMART; SM00320; WD40; 1.
DR SMART; SM00678; WD_REPEATS; UNKNOWN_6.
KW Ligase; Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

      | | | | | : : : : | | | | | : | | | | : | | | | :
Query Match 20.2%; Score 635; DB 11; Length 569;
Best Local Similarity 30.3%; Pred. No. 1.3e-45;
Matches 164; Conservative 92; Mismatches 214; Indels 72; Gaps 14;

Qy 53 KLDHGEVRSFSLGKKPKCV-----SEYTSITGLV-----PCSATPTTFGDL 94
      | | | | | : | | | | : | | | | : | | | | :
Db 14 KFMNSSEDCNCGEPKPKIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTCNCVAKA 73

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 95 RAANGOG-----QRRRTTSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKH 149
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 74 KLANGTSSMIVPKQKLSASYEKEKEL--CVKYFEQWSESDQVEFVHLISQMCHYQGHG 131

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 150 MMQVTEPQFQDFISLSP-----KELALYVLSFLEPKDLLQAAQTCRYWRIIAEDNLLWRE 205
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 132 INSYLKPMLQDFITALPARGLDHIAENILSYLDKASLCAELVCKEYRVTSQGLMKWK 191

      | | | | | : : : : | | | | | : | | | | : | | | | :
Qy 206 KCKE-----EGIDEPLHKRRK-----VIRPGFIHSPKSAI-----IRQ--HRI 243
      | | | | | : : : : | | | | | : | | | | : | | | | :
Db 206 KCKE-----EGIDEPLHKRRK-----VIRPGFIHSPKSAI-----IRQ--HRI 243
```

```
Db 192 LIERMVRTDLSLWRLAE-----RRGWGOYLFKNKPPDENAPPNSFYRALYPKIIQDIETI 246
QY 244 DYNWRGELKSPKV-LKGHDDHVITCLOFCGNRIYVSGSDNTLKVWSAVTGKCLRTLUGH 302
Db 247 ESNWRCGRISLQRIHCRSETSGVYCLQYDDQKIVSGLRDNITIKIWDKSTLECKRILTGH 306
QY 303 TGGVSSQMRDNIISGSTDRTLKVNNAETGCIHTLYGHTSTVRCMHILHEKRVVSGSRD 362
Db 307 TGSVLCLOQYDERVITGSSDSTVRVMDVNAEMNLTHHCEAVLHLRFNNGMMVTCSD 366
QY 363 ATLRLVMDIETGQCL---HYLMGHVAAVRCVQYDGRVVSAGYDFVWKVWDPEPETETCLHTL 419
Db 367 RSTAVWDMASPTDITLRLRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEVFTL 426
QY 420 QGHTNRVYSIQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHOSLTSGMELKNLTVSG 479
Db 427 NGHKGACIACQYDRDLVWSGSDNTIRLWDIECGACLRVLEGEELVRCIRFNDKRVISG 486
QY 480 NADSTVKIWDIKTG-----QCLOTLOGPNKHQSAYTCLQFNKNFVITSSDDGTVKL 530
Db 487 AYDGKIKVWDLMAALDPRAPAGTLCRLTLV---EHSGRVFRLOQDFEQIVSSSHDDTILI 543
QY 531 WD 532
Db 544 WD 545

RESULT 7
QYR167
ID Q9RIG7 PRELIMINARY; PRT; 569 AA.
AC Q9RIG7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TReMBLrel. 16, Last annotation update)
DE BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF110396; AAD41025.1; -.
DR InterPro; IPR001680; -.
DR InterPro; IPR001810; -.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS0181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_6.
DR SMART; SM00320; WD40; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Query Match 20.0%; Score 630; DB 11; Length 569;
Best Local Similarity 30.1%; Pred. No. 3.6e-45;
Matches 163; Conservative 91; Mismatches 216; Indels 72; Gaps 14;

QY 53 KLDHGEVRSFSLGKKPKCV-----SEYTSITGLV-----PCSAITPTFGDL 94
Db 14 KPMNSEREDCNGEPKRIIPEKNSLRQTYNSCARLCINQETVCLTSTAMKTCNCVAKA 73
QY 95 RAANGSQG-----QQRRTTSVQPPPTGLQEWLKMFSQSGPEKLLALDELIDSCETQVKH 149
Db 74 KLANGTSSWIVPKQRLKLSAYSEKEKL--CVKYFFQWSESQVFEVHLSIQMCHYQGH 131
QY 150 MMOVTEPQFQDFISLPL-----KELALYVLFLEPKDLQAAQTCRYWRILAEDNLLWRE 205
Db 132 INSYLKPLQDFITAPLARGLDHTENILSYLDAKSILCAAEVLCKEYRYVTSQGLWKK 191
QY 206 KCKE-----EGIDEPLHKKRK-----VIRPGFIHSPWKSAY-----IRO--HRI 243
```

```
Db 192 LIERMVRTDLSLWRLAE-----RRGWGOYLFKNKPPDENAPPNSFYRALYPKIIQDIETI 246
QY 244 DYNWRGELKSPKV-LKGHDDHVITCLOFCGNRIYVSGSDNTLKVWSAVTGKCLRTLUGH 302
Db 247 ESNWRCGRISLQRIHCRSETSGVYCLQYDDQKIVSGLRDNITIKIWDKSTLECKRILTGH 306
QY 303 TGGVSSQMRDNIISGSTDRTLKVNNAETGCIHTLYGHTSTVRCMHILHEKRVVSGSRD 362
Db 307 TGSVLCLOQYDERVITGSSDSTVRVMDVNAEMNLTHHCEAVLHLRFNNGMMVTCSD 366
QY 363 ATLRLVMDIETGQCL---HYLMGHVAAVRCVQYDGRVVSAGYDFVWKVWDPEPETETCLHTL 419
Db 367 RSTAVWDMASPTDITLRLRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEVFTL 426
QY 420 QGHTNRVYSIQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHOSLTSGMELKNLTVSG 479
Db 427 NGHKGACIACQYDRDLVWSGSDNTIRLWDIECGACLRVLEGEELVRCIRFNDKRVISG 486
QY 480 NADSTVKIWDIKTG-----QCLOTLOGPNKHQSAYTCLQFNKNFVITSSDDGTVKL 530
Db 487 AYDGKIKVWDLMAALDPRAPAGTLCRLTLV---EHSGRVFRLOQDFEQIVSSSHDDTILI 543
QY 531 WD 532
Db 544 WD 545

RESULT 8
QYR167
ID Q44382 PRELIMINARY; PRT; 510 AA.
AC Q44382;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE SLIMB.
OS SLMB OR SLIMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98121115; PubMed=9461217;
RA Jiang J., Struhl G.;
RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-
box/WD40-repeat protein Slmb.";
RL Nature 391:493-496(1998).
DR EMBL; AF032878; AAC38852.1; -.
DR FlyBase; FBgn0023423; slmb.
DR InterPro; IPR001680; -.
DR InterPro; IPR001810; -.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR SMART; SM00320; WD40; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 19.6%; Score 617; DB 5; Length 510;
Best Local Similarity 30.6%; Pred. No. 3.9e-44;
Matches 146; Conservative 91; Mismatches 190; Indels 50; Gaps 10;

QY 104 RRRITSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCETQVKHMMQVIEPQFQDFI 163
Db 30 RKKDSSTPYOTRELCFQVFTQWSESGQVDFVEHLLSRMCHYQHGINLAYLKPMLQRDFI 89
QY 164 SLLP-----KELALYVLFLEPKDLQAAQTCRYWRILAEDNLLWRE---KCKEGIDEP 215
Db 90 TLLPLKGLDHGENTILSYLDAESKSELVCKEELRVISSEGMLWKKLIERKVRTDLSLWRLG 149
```

Qy 216 LHKRRKVIKPGFIHSPWKSAYIRQH-----RIDTNRRGELKSPKV-LK 259  
 Db 150 L-AERRNMWQ--YLFKPRPGQTQPHSFHRELFPKIMNDISINNRTGRHMLRRINCR 206  
 Qy 260 GHDDHVITCLOFCGCRNRIYVSGDDNTLKWMSAVTGKCLRTLVTGHTGGVSSQMRDNIISG 319  
 Db 207 SENSKGVCVCLQYDDGKIVSGULRDNTIKIWDRTDLQCVRKTLGHTGVSVCVLCQYDDKVIISG 266  
 Qy 320 STDRTLKVNNAETGECITHTLYGHTSTVRCMHLHEKRVVSGSRDATLRWDIETGQCL- 376  
 Db 267 SSDSTVRVMDVNTGEMVNTLIHCEAVLHLFRNNGMMVTCSKDRSIAMVDMTSPSEITLR 326  
 Qy 377 HVLGMHVAARVCYDGRVVRVSGAYDFMVKVMDPETETCLTHTLOGHTNRVYSLOFDGIHV 436  
 Db 327 RVLVGHRAAVNVDFDEKIVIVSAGDRTIKVMTSTCFEFTVTLNGHKGACIACQYDRDLV 386  
 Qy 437 VSGSLDTSIRVMDVETGNCIHTLGHOSLTSGMELKONILVSGNADSTVKIWDI----- 490  
 Db 387 VSGSSDINSIRLWDIECGACLRVLSGHEELVRCIRFDTKRIVSGAYDGKIKVMDLVAALDP 446  
 Qy 491 ---KTGQCLQTLQGNPKHQSAVTCLOQFNKNVITSSDGTGTVKLWDL-----KTG 536  
 Db 447 RAASNTLCLNTLV---EHTGRVFRLOQDFEIVSSSHDDTLIWDLFNFTPNENKGTG 500

RESULT 9  
 Q9VDE3 ID Q9VDE3 PRELIMINARY; PRT; 510 AA.  
 AC Q9VDE3;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE SLMB PROTEIN (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=BERKELEY.  
 RX MDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobery C.C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "the genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVARY, AND IMAGINAL DISC;  
 RX MEDLINE=20245299; PubMed=10781936;  
 RA Miletich I., Limbourg-Bouchon B.;  
 RT "Drosophila null limb clones transiently deregulate Hedgehog-  
 RT independent transcription of wingless in all limb discs, and induce  
 RT decapaplegic transcription linked to imaginal disc regeneration.";  
 RL Mech. Dev. 93:15-26(2000).  
 DR EMBL; AE003733; AAF55853.1; -;  
 DR EMBL; AF222924; AAF63214.1; -;  
 DR EMBL; AF222923; AAF63213.1; -;  
 DR FlyBase; FBgn0023423; slmb.  
 DR InterPro; IPR001680; -;  
 DR Pfam; PF00400; WD40; 7.  
 DR Pfam; PF00646; F-box; 1.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR PROSITE; PS00181; FBOX; 1.  
 DR PROSITE; PS00678; WD\_REPEATS; 5.  
 DR SMART; SM00256; FBOX; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match 19.6%; Score 617; DB 5; Length 510;  
 Best Local Similarity 30.68; Pred. No. 3.9e-44;  
 Matches 146; Conservative 92; Mismatches 189; Indels 50; Gaps 10;

Qy 104 RRRITSVQPTGLGELWLMFQSGPEKLLALDELIDSCPTQVKMMQVIEPQQRDFI 163  
 Db 30 RKDSSPTYQTERELCFQYFTQWSEGVDFVHLLSRMCHYQHGINAYLKPMLQRDFI 89  
 Qy 164 SLLP----KELALYVLFLEPKDLLQAAQCYRYWILAEADNLLWRE----KCKEGIDEP 215  
 Db 90 TLLPIKGLDHTAENILSYLDAESLKSSELVCKEVLRISEGLMWMKLFIERKVRVDSLWRG 149  
 Qy 216 LHKRRKVIKPGFIHSPWKSAYIRQH-----RIDTNRRGELKSPKV-LK 259  
 Db 150 L-AERRNMWQ--YLFKPRPGQTQPHSFHRELFPKIMNDISINNRTGRHMLRRINCR 206  
 Qy 260 GHDDHVITCLOFCGCRNRIYVSGDDNTLKWMSAVTGKCLRTLVTGHTGGVSSQMRDNIISG 319  
 Db 207 SENSKGVCVCLQYDDGKIVSGULRDNTIKIWDRTDLQCVRKTLGHTGVSVCVLCQYDDKVIISG 266  
 Qy 320 STDRTLKVNNAETGECITHTLYGHTSTVRCMHLHEKRVVSGSRDATLRWDIETGQCL- 376  
 Db 267 SSDSTVRVMDVNTGEMVNTLIHCEAVLHLFRNNGMMVTCSKDRSIAMVDMTSPSEITLR 326  
 Qy 377 HVLGMHVAARVCYDGRVVRVSGAYDFMVKVMDPETETCLTHTLOGHTNRVYSLOFDGIHV 436  
 Db 327 RVLVGHRAAVNVDFDEKIVIVSAGDRTIKVMTSTCFEFTVTLNGHKGACIACQYDRDLV 386  
 Qy 437 VSGSLDTSIRVMDVETGNCIHTLGHOSLTSGMELKONILVSGNADSTVKIWDI----- 490  
 Db 387 VSGSSDINSIRLWDIECGACLRVLSGHEELVRCIRFDTKRIVSGAYDGKIKVMDLVAALDP 446  
 Qy 491 ---KTGQCLQTLQGNPKHQSAVTCLOQFNKNVITSSDGTGTVKLWDL-----KTG 536  
 Db 447 RAASNTLCLNTLV---EHTGRVFRLOQDFEIVSSSHDDTLIWDLFNFTPNENKGTG 500

RESULT 10  
 Q9P7V1



[illegible]

```
Db 718 FTCKIFSLTKTKTLTYTNHQAENVTCINYLGDVENKNCITSSLDKTIQLWDAETGSCISLTL 777
Qy 420 OGHTNRVYSLQFDGI-----H-VVSGSLDTSIRVWDVETGNCIHTLTGHQSILTSME 470
Db 778 RGHGGIYCVTDQVATHGNGYNHLVVSADYKTSNWDTRSSSKVRSFTOHTEDVLCCY 837
Qy 471 LKDNILVSGNADSTVKIWDIKTGQCLQF--LOGPNKHQSAVTCLOFNKFNVTSSDDGTVK 529
Db 838 VFDQVVTGSCDGIKLDWIDGTGKTISTIFSEYRQKNYVTVQDFDQSKIISSKGTGIIR 897
Qy 530 LWDLTGTFIRNLVLESGGGVVWRIRASNTKLV 565
Db 898 IWDIYNERDSRSI-----GGHETIFSLQFNQKLI 928

RESULT 12
Q9NMW7
ID Q9NMW7 PRELIMINARY; PRT; 334 AA.
AC Q9NMW7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CDNA FLJ20545 FIS, CLONE KAT11476.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isoigai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/genBank/DBJ databases.
DR EMBL; AK000552; BAA91248.1; -.
DR InterPro; IPR001632; -.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00319; GPROTEINB.
DR PROSITE; PS00320; GPROTEINRPT.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_4.
KW Repeat; WD repeat.
SQ SEQUENCE 334 AA; 36588 MW; 4BF30914A2250286 CRC64;

Query Match 14.7%; Score 461; DB 4; Length 334;
Best Local Similarity 33.9%; Pred. No. 4e-31;
Matches 102; Conservative 68; Mismatches 93; Indels 38; Gaps 13;

Qy 258 LKGHDDHVTCLQFC--GNRIYSGDDNTLKVSAVTKGKCLRTLVTGHTGGV----WSSQM 311
Db 41 LAGH-TKAVSSVKFSPNGEWLASSADKLKINGAYDGKFEKTLISGHKLGISDVANSSD- 98
Qy 312 RDNIIISGSDTORTLKVNAETGECIHTLYGHTSTVRCMHLHEKR--VVSGRSDATLRWD 369
Db 99 -SNLIVSADDTKLKIMDVSSGCKLTKLGHSNVYVFCNFPQSNLIVSGSFDESRIWD 157
Qy 370 IETGQCLHVLGHVNAVRCVQY--DGRVVSAGVDFWVKVWDPEETCLHTLQGHNTNRVY 427
Db 158 VTKGKCLTLPAHSDPPVSAVHFNRDGLSIVSSYDGLCRWDIASGQCKLTLDDNDPPV 217
Qy 428 S-LQF--DGIHVSGSLDTSIRVWDVETGNCIHTLTGHQ-----SLTSGMELKDN 474
Db 218 SFVRFSPNGKYILAATLNDTLKLDYSGKGLKTYTGKNEKYCIFANFSVTG-----K 272
Qy 475 ILVSGNADSTVKIWDIKTGQCLQFLOGPNKHQSAV--TCLQFNKNFVITSS--DDGTVKL 530
Db 273 WIVSGSDNLVYIWNLTQKEIVQKLOG---HTDVIIVISTACHPTENIATSAALENDKTIKL 329
Qy 531 W 531
Db 330 W 330
```

```
RESULT 13
Q9PTR5
ID Q9PTR5 PRELIMINARY; PRT; 410 AA.
AC Q9PTR5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Lisl.
GN Lisl.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Reiner O., Shmueli O.;
RT "Characterization of the chicken homolog of Lisl.";
RL Submitted (DEC-1998) to the EMBL/genBank/DBJ databases.
DR EMBL; AF113946; AAF18938.1; -.
DR InterPro; IPR001680; -.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_5.
DR SMART; SM00320; WD40; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 410 AA; 46664 MW; FC5848D06E0DCA20 CRC64;

Query Match 14.0%; Score 439; DB 13; Length 410;
Best Local Similarity 29.9%; Pred. No. 4e-29;
Matches 109; Conservative 63; Mismatches 133; Indels 60; Gaps 10;

Qy 227 GFHSPPKSAVIRQHR-----DTNRRRGELKSPK-----VLKGDH 262
Db 49 GLEKKWTSVIRLQKKVMELESKLEKEETSGGPLQKQKRPKEWIPPEKYALSGRH 108
Qy 263 DHVITCL-QFCGNRIYSGDDNTLKVSAVTKGKCLRTLVTGHTGGV--SSQMRDNIISG 319
Db 109 SPVTRVIFHPVFSVMVSASEDATIKVWDYETGDFERTLKGTDSVQDISFDHTGKLASC 168
Qy 320 STDRTLKVNAETGECIHTLYGHTSTVRCMHL--HEKRVVSGSRDATLRVWDIETGQCLH 377
Db 169 SADMTIKLWDFQGFECIRTMHGHNVSSVAIMPNGDHIVSASRDKTIKMWEEVQYGVK 228
Qy 378 VLMGHVAARVCQ--YDGRVVSAGYDFWVKVWDPEETCLHTLQGHNTNRVYSLQF---- 431
Db 229 TFGHREVRVVRPNQDGTLIASCNDQTVRVVVVATKECKAELEHEHVVVECTISWAPES 288
Qy 432 -----DGIHVSGSLDTSIRVWDVETGNCIHTLTGHQSILTSMEKLD 473
Db 289 SYSTISEATGSETKSKGKPGFLLSGSRDKTIKMWIDISTGCLMTLVGHDNVVRGVLFHS 348
Qy 474 --NILVSGNADSTVKIWDIKTGQCLQFLOGPNKHQSAVTCLOFNKN--FVITSSDDGTVK 529
Db 349 GKGKILSCADDTLRVWDFKNKCMKTL--NAHEHFTSLDFHKTAPYVTVGSVDQIVK 405
Qy 530 LWDLK 534
Db 406 WVECR 410

RESULT 14
Q9GL51
ID Q9GL51 PRELIMINARY; PRT; 410 AA.
AC Q9GL51;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE IB-ALPHA SUBUNIT.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



---